



# Fusion de données multi-capteurs pour l'habitat intelligent

Damien Brulin

## ► To cite this version:

Damien Brulin. Fusion de données multi-capteurs pour l'habitat intelligent. Autre. Université d'Orléans, 2010. Français. NNT : 2010ORLE2066 . tel-01236617

**HAL Id: tel-01236617**

**<https://theses.hal.science/tel-01236617>**

Submitted on 2 Dec 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**UNIVERSITÉ D'ORLÉANS**



**ÉCOLE DOCTORALE SCIENCES ET TECHNOLOGIES**

Institut PRISME

**THÈSE** présentée par :

**Damien BRULIN**

soutenue le : **27 Août 2010**

pour obtenir le grade de : **Docteur de l'université d'Orléans**

Discipline/ Spécialité : **Sciences et Technologies Industrielles**

**Fusion de données multi-capteurs pour l'habitat intelligent**

**THÈSE dirigée par :**

**Youssef TOURÉ**

Professeur des Universités, Université d'Orléans

**RAPPORTEURS :**

**Laurent FOULLOY**

Professeur des Universités, Université Chambéry

**Norbert NOURY**

Professeur des Universités, Université Lyon 1

**JURY :**

**Jean-Yves CADOREL**

Ingénieur, CRESITT Industries

**Estelle COURTIAL**

Maître de Conférences, Université d'Orléans

**Laurent FOULLOY**

Professeur des Universités, Université Chambéry

**Thierry-Marie GUERRA**

Professeur des Universités, Université Valen-  
ciennes, Président du jury

**Norbert NOURY**

Professeur des Universités, Université Lyon 1

**Youssef TOURÉ**

Professeur des Universités, Université d'Orléans







# Remerciements

C'est une grande étape qui s'achève, avec une ligne d'arrivée symbolisée par ce manuscrit de thèse. Quatre années de travail, de rencontres et d'échanges, plus ou moins dans la douleur, mais toujours dans la joie et la bonne humeur. J'ai hésité un moment avant de me lancer dans cette aventure mais la perspective de travailler sur un sujet pluridisciplinaire et avec une équipe composée de trois autres "braves" m'ont convaincu de faire le grand saut.

Ce travail de thèse n'aurait pu être possible sans une personne d'exception, Madame Estelle Courtial, Maître de Conférences à l'Université d'Orléans. Elle a su faire preuve de patience, d'enthousiasme et d'une grande disponibilité pour me guider et me soutenir de près comme de loin. Il n'est jamais facile de prendre en charge l'encadrement d'une thèse en cours de route, mais en s'appuyant sur sa rigueur et ses compétences scientifiques, elle a su mener à bon port un bateau qui naviguait en eaux troubles. Je suis fier des travaux accomplis et c'est en grande partie grâce à toi. Merci pour ta sympathie et ton soutien Estelle.

Je remercie Monsieur Youssef Touré, Professeurs des Universités de l'Université d'Orléans, pour avoir accepté de diriger cette thèse et de m'avoir accueilli au sein de son laboratoire. Ce travail de thèse s'inscrivant au sein du projet CAPTHOM du pôle de compétitivité S2E2, je remercie également Monsieur Thierry Allard, Président du pôle S2E2, d'avoir permis la réalisation de ces travaux. Je remercie enfin Monsieur Joël Allain et Madame Anne-Marie Jolly, respectivement Directeur de l'Ecole Nationale Supérieure d'Ingénieurs de Bourges et Directrice de Polytech'Orléans, de m'avoir accueilli successivement au sein de leur établissement et de m'avoir permis de mener cette thèse dans les meilleures conditions possibles.

Je remercie Messieurs Laurent Foulloy et Norbert Noury, respectivement Professeur des Universités à Polytech'Savoie de l'Université de Chambéry et Professeur des Universités à l'Université Lyon 1, qui ont accepté la charge de rapporteur de ces travaux de

thèse. Je les remercie également de leurs présences au sein de mon jury, qu'ils trouvent ici l'expression de ma plus profonde reconnaissance.

Je remercie également l'ensemble des membres du jury qui ont accepté d'examiner mes travaux avec la plus grande attention et de s'être rendu disponible peu de temps avant la rentrée universitaire : Monsieur Jean-Yves Cadorel, Responsable Recherche et Développement chez CRESITT Industrie et Monsieur Thierry-Marie Guerra, Professeur des Universités à l'Université de Valenciennes et du Hainaut Cambrésis. Qu'ils trouvent ici l'expression de mes sentiments les plus sincères.

J'aimerais adresser mes remerciements à l'ensemble des partenaires industriels du projet CAPTHOM pour leur écoute, leurs conseils et l'intérêt qu'ils ont porté à l'ensemble de mes travaux. Un grand remerciement à Monsieur Yves Parmentier, animateur du pôle Capteurs et Automatismes, véritable chef d'orchestre de ce projet pour sa grande disponibilité et son investissement afin que chacun puisse travailler sereinement. Une pensée également à Madame Laure Spina pour sa gentillesse et son sérieux.

Mes remerciements chaleureux à Monsieur Frédéric Kratz, Professeur des Universités à l'ENSIB, pour m'avoir proposé ce travail de recherche et pour m'avoir accueilli au sein de l'équipe MCDS. Une pensée particulière pour Monsieur Vincent Idasiak, alias Joe Dariak. Je te remercie pour ton soutien de tous les jours, de l'opportunité que tu m'as donnée d'enseigner avec chaque année davantage de responsabilité (les murs de l'ENSIB porteront à jamais la marque de cette preuve de confiance) et bien entendu pour ton humour. Je souhaite également exprimer toute ma reconnaissance aux occupants des bureaux voisins à l'ENSIB qui m'ont permis de travailler dans une ambiance conviviale. Merci donc à Hélène, Serge, Adel, Benoît et Ludovic.

J'aimerais également adresser mes plus chaleureux remerciements à l'ensemble des doctorants que j'ai croisés au cours de ces quatre années. Je pense tout particulièrement à Hazem, Hamid, David, Rudy, Baptiste et Robin. Un grand merci à Guillaume pour ses conseils et son enthousiasme et à Jérôme, mon cousin, pour son écoute et son soutien dans les moments les plus pénibles. Une pensée toute particulière à Antoine, pour ta gentillesse et ton courage, à Yannick pour ton amitié. Enfin un énorme merci à Pierrot pour ses trois années passées en ta compagnie dans notre bureau de l'ENSIB, trois années de partage, de soutien mutuel et d'amitié. Merci pour tout.

Je terminerai ces remerciements par une pensée à mes amis qui ont suivi de loin mon

parcours mais qui ont toujours été présents. Merci donc à Thomas, Benjamin, Christophe, Julien, Ella, Vanessa et Sébastien. Une pensée émue à toute ma famille pour leur soutien inconditionnel et leur amour, en particulier mes parents qui m'ont toujours suivi dans mes choix et m'ont permis de réaliser ce parcours. Merci à mes grands mères et à mon grand père qui malheureusement n'a pas pu voir l'aboutissement de ces travaux, mais dont il aurait été très fier. Merci à mes beaux-parents et notamment à Halina dont la présence lors de ma soutenance m'a énormément touché. Enfin, un immense merci à toi, Catherine, pour ton amour de tous les jours , pour ton écoute attentive et ton soutien de tous les instants, toi qui a rendu possible l'ensemble de ces travaux et qui m'a permis d'être ce que je suis aujourd'hui. Grâce à toi j'ai trouvé un équilibre et ce travail s'est également le tien. Merci du plus profond de mon coeur.





# Table des matières

<b>Introduction</b>	<b>13</b>
<b>1 La détection de présence humaine</b>	<b>17</b>
1.1 Le projet CAPTHOM . . . . .	18
1.1.1 Le contexte du projet . . . . .	18
1.1.2 Les enjeux de la détection de présence . . . . .	21
1.1.2.1 Les enjeux pour CAPTHOM . . . . .	21
1.1.2.2 Le contexte national et international . . . . .	22
1.2 La détection de présence humaine . . . . .	25
1.2.1 La problématique de détection de présence humaine . . . . .	25
1.2.2 Les solutions technologiques et leurs limitations . . . . .	27
1.2.2.1 Les technologies basées sur le rayonnement infrarouge . . .	27
1.2.2.2 La caméra vidéo . . . . .	32
1.2.2.3 Les autres technologies servant à la détection de présence .	34
1.2.3 L'homme, centre du processus de détection de présence . . . . .	36
1.3 Conclusion . . . . .	37
<b>2 La localisation 3D par estimation visuelle sur horizon glissant</b>	<b>41</b>
2.1 L'estimation de position en traitement d'images . . . . .	42
2.1.1 Les approches de localisation 3D en traitement d'images . . . . .	42
2.1.2 L'estimation de position par filtrage de Kalman . . . . .	45
2.2 L'estimation de position sur horizon glissant . . . . .	47
2.2.1 Le principe de l'estimation d'état sur horizon glissant . . . . .	47
2.2.2 L'estimation visuelle sur horizon glissant . . . . .	50
2.2.2.1 Le modèle global . . . . .	50
2.2.2.2 La formulation mathématique du VRHE . . . . .	57
2.3 Application du VRHE à la localisation 3D d'une personne . . . . .	59
2.3.1 Les conditions expérimentales . . . . .	59
2.3.2 Expérience sans contrainte . . . . .	63

2.3.2.1	Comparaison entre le modèle déplacement linéaire et le modèle non holonome . . . . .	64
2.3.2.2	Influence de l'horizon d'estimation $N_e$ . . . . .	65
2.3.2.3	Comparaison avec le filtre de Kalman étendu (EKF) . . . .	70
2.3.3	Expérience avec contrainte . . . . .	71
2.4	Conclusion . . . . .	75
<b>3</b>	<b>La fusion de données multi-capteurs</b>	<b>77</b>
3.1	La fusion de données pour le projet CAPTHOM . . . . .	78
3.1.1	Pourquoi la fusion de données dans CAPTHOM? . . . . .	78
3.1.1.1	L'obtention d'informations pertinentes . . . . .	78
3.1.1.2	La prise de décision . . . . .	79
3.1.2	Les données à fusionner . . . . .	82
3.2	Généralités en fusion de données . . . . .	83
3.2.1	Définition de la fusion de données . . . . .	84
3.2.2	Les données, élément essentiel de la fusion . . . . .	86
3.2.2.1	Les caractéristiques des données . . . . .	86
3.2.2.2	Les imperfections des données . . . . .	87
3.2.3	Modélisation du processus de fusion de données . . . . .	89
3.2.3.1	Les niveaux de fusion . . . . .	89
3.2.3.2	Les architectures de fusion . . . . .	93
3.2.4	Les théories et méthodes utilisées en fusion . . . . .	94
3.2.4.1	Les approches probabilistes et numériques . . . . .	95
3.2.4.2	Les approches symboliques . . . . .	96
3.2.5	Conclusion sur la fusion de données . . . . .	101
3.3	Architecture de fusion pour le projet CAPTHOM . . . . .	103
3.3.1	Fusion pour la détection de présence humaine . . . . .	104
3.3.1.1	Processus de détection de présence humaine . . . . .	104
3.3.1.2	Exemples d'application . . . . .	105
3.3.2	La localisation 2D et 3D . . . . .	112
3.3.2.1	La localisation 3D par zone . . . . .	113
3.3.2.2	La localisation 3D précise . . . . .	117
3.3.3	Fusion pour la détermination de posture . . . . .	119
3.3.3.1	Méthodologie . . . . .	121
3.3.3.2	Résultats expérimentaux . . . . .	128
3.3.4	Fusion pour la prise de décision . . . . .	129
3.4	Conclusion . . . . .	138

<b>Conclusion</b>	<b>140</b>
<b>Publications de l’auteur</b>	<b>144</b>
<b>Références</b>	<b>146</b>
<b>Annexes</b>	<b>156</b>



# Table des figures

1.1	Représentation d'un capteur pyroélectrique. . . . .	28
1.2	Lobes de détection d'un détecteur infrarouge passif. . . . .	30
1.3	Exemple de détecteurs infrarouges passifs commercialisés. . . . .	30
1.4	Schéma de l'effet Seebeck. . . . .	31
2.1	Principe de l'estimation d'état au temps $k.T_e$ . . . . .	49
2.2	Principe de l'estimation d'état au temps $(k + 1).T_e$ . . . . .	49
2.3	Représentation des repères définis lors de la calibration de la caméra. . . . .	51
2.4	Plan image. . . . .	53
2.5	Estimation de position à l'aide d'une caméra perspective. . . . .	56
2.6	Environnement vu par la caméra perspective. . . . .	60
2.7	Les différentes phases de l'algorithme de traitement d'images. . . . .	61
2.8	Représentation du repère $R_h$ utilisé pour l'homographie. . . . .	63
2.9	Comparaison entre les modèles linéaire et non holonome suivant l'axe x dans le repère $R_h$ . . . . .	64
2.10	Comparaison entre les modèles linéaire et non holonome suivant l'axe y dans le repère $R_h$ . . . . .	64
2.11	Représentation du trajet N°1. . . . .	65
2.12	Estimation du trajet 1 pour $N_e = 5, 10, 15$ . . . . .	66
2.13	Les positions en x et y dans $R_h$ pour le trajet N°1 avec $N_e = 5, 10, 15$ . . . . .	66
2.14	Erreur pour le trajet N°1 dans le repère $R_h$ avec $N_e = 5, 10, 15$ . . . . .	67
2.15	Représentation du trajet N°2. . . . .	67
2.16	Estimation du trajet 2 pour $N_e = 5, 10, 15$ . . . . .	68
2.17	Les positions en x et y dans $R_h$ pour le trajet N°2 avec $N_e = 5, 10, 15$ . . . . .	68
2.18	Erreur pour le trajet N°2 dans le repère $R_h$ avec $N_e = 5, 10, 15$ . . . . .	69
2.19	Comparaison des estimations de l'EKF et du VRHE suivant les axes x et y du repère $R_h$ . . . . .	70
2.20	Influence d'un obstacle sur le résultat de détection de présence. . . . .	71
2.21	Influence des occultations sur la mesure de la position suivant $y_h$ . . . . .	72

2.22	Exemples de prise en compte de contrainte sur le déplacement sur deux trajets. . . . .	73
2.23	Prise en compte des variations lumineuses. . . . .	74
3.1	Objectifs du maintien à domicile et de la gestion énergétique. . . . .	79
3.2	Le dispositif CAPTHOM. . . . .	82
3.3	Modèle de la fusion de données défini par le JDL [Hal09]. . . . .	90
3.4	Représentation des processus de fusion de niveau 0 et 1 [Cro93]. . . . .	91
3.5	Révision du modèle du JDL [Bow04]. . . . .	92
3.6	Représentation de la fusion de mesures. . . . .	93
3.7	Représentation de la fusion de primitives. . . . .	94
3.8	Représentation de la fusion de décisions. . . . .	94
3.9	Schéma de l'architecture de fusion de données. . . . .	103
3.10	Organigramme de la détection de présence humaine. . . . .	105
3.11	Placement et couverture des capteurs pour le scénario N°1. . . . .	106
3.12	Trajet suivi par la personne lors du scénario N°1. . . . .	107
3.13	Réponse des capteurs et décision finale pour le scénario N°1. . . . .	108
3.14	Réponse des capteurs et décision finale pour le scénario N°2. . . . .	109
3.15	Réponse des capteurs et décision finale pour le scénario N°2 avec mémorisation. . . . .	110
3.16	Réponse des capteurs et décision finale pour le scénario N°3 avec mémorisation. . . . .	111
3.17	Organigramme de détermination de la localisation. . . . .	112
3.18	Découpage en zones de l'appartement témoin du GIS MADONAH. . . . .	113
3.19	Placement des détecteurs IRP et exemples de scénarios simulés. . . . .	114
3.20	Résultat obtenu pour le scénario du tableau 3.4. . . . .	115
3.21	Résultat obtenu pour un second scénario. . . . .	117
3.22	Placement et couverture des capteurs pour la localisation. . . . .	118
3.23	Trajectoire réelle et trajectoire estimée par la caméra. . . . .	119
3.24	Organigramme de la détermination de la posture. . . . .	120
3.25	Représentation du point moyen et de l'axe principal du nuage de points. . . . .	121
3.26	Représentation des distances $D1$ , $D2$ et $D3$ . . . . .	122
3.27	Variations de $r_1$ suivant deux scénarios. . . . .	124
3.28	Variations de $r_2$ suivant deux scénarios. . . . .	124
3.29	Fonctions d'appartenance des entrées $r_2$ et $ax_1$ . . . . .	124
3.30	Illustration de la règle N°1. . . . .	126
3.31	Fonction d'appartenance globale du système. . . . .	127

3.32	Exemple de postures utilisées pour les tests. . . . .	128
3.33	Prise de décision pour la gestion énergétique dans l'habitat. . . . .	130
3.34	Fonctions d'appartenance du système d'éclairage. . . . .	131
3.35	Prise de décision pour le maintien à domicile. . . . .	132
3.36	Situation de vigilance. . . . .	133
3.37	Scénario de chute. . . . .	134
3.38	Non déclenchement du dispositif d'éclairage de la pièce. . . . .	135
3.39	Déclenchement du dispositif d'éclairage en fin de journée. . . . .	136
3.40	Absence de présence humaine. . . . .	137
3.41	Correspondance entre les points pour le calcul d'homographie. . . . .	157





# Introduction

Depuis ces dernières années, le concept d’habitat intelligent s’est largement développé aussi bien pour la gestion énergétique de bâtiments ou de maisons individuelles que pour le maintien à domicile de personnes âgées ou handicapées. La gestion de l’énergie est actuellement une préoccupation majeure : économiser l’énergie, maîtriser sa consommation, la diversifier sont les différents arguments pour des applications domotiques. D’autre part, face au vieillissement de la population et compte-tenu du nombre insuffisant de places dans les établissements spécialisés (ces derniers n’ont pas pu ou su s’adapter aux conséquences de l’augmentation de l’espérance de vie), le maintien à domicile des personnes âgées apparaît comme une solution alternative.

De nombreux projets académiques et industriels, s’intéressent au problème du confort et de la gestion d’énergie sous la forme d’habitat communicant [Kid99][Mas99]. On peut citer entre autres, le projet E2Home entre Ericson et Electrolux [Yeo08] ou encore le projet Homelab de Philips. Au niveau national, la société Legrand a conçu une maison domotique, miroir des nouvelles technologies développées [Leg05].

Concernant les habitats intelligents pour la santé, il existe de nombreux projets allant du maintien à domicile (PROSAFE [Cha03][Bon08b]) à la télé-surveillance médicale (EMUTEM [Med09], TelePat [Bou06], HIS [Duc04], AILISA [Nou05]).

Un grand nombre de capteurs (capteurs infrarouges, accéléromètres, capteurs de pression sanguine, microphones,...) sont nécessaires pour informer de la situation, normale ou anormale.

Ces différents enjeux ont en commun une problématique qui est la détection de présence humaine. C’est dans ce contexte que s’inscrit ce travail de recherche, supporté par un projet industriel, le projet CAPTHOM. L’objectif du projet CAPTHOM est de développer un dispositif ou capteur intelligent de détection de présence humaine pour la gestion énergétique dans l’habitat et le maintien à domicile de personnes âgées. Ce dispositif doit permettre :

- de détecter la présence d’une personne ;

- de localiser cette personne dans l'environnement ;
- de compter le nombre de personnes présentes (pour la gestion de l'énergie) ;
- de déterminer la posture de la personne (pour le maintien à domicile).

A cela, s'ajoutent des contraintes et des exigences telles que la non-invasivité du dispositif (pas de capteur embarqué), l'utilisation de technologies existantes et la robustesse du dispositif face aux perturbations (fausses détections).

Nous allons proposer un dispositif permettant de répondre à ces exigences.

- On va montrer que la détection de présence humaine nécessite plusieurs capteurs et que ces différentes informations, entachées de bruits et de perturbations, devront être fusionnées pour garantir une information fiable en sortie du dispositif.
- La localisation 3D d'une personne à partir d'informations visuelles sera abordée sous forme de problème d'optimisation non linéaire dans le plan image.
- La détermination de la posture sera étudiée sans capteur embarqué sur la personne et développée grâce à un module de fusion de données.
- Enfin, on proposera une architecture générale de fusion pour le dispositif CAPTHOM, allant de la détection de la personne jusqu'à la prise de décision.

Ce mémoire est organisé en trois chapitres structurés de la manière suivante :

- Dans le premier chapitre, le contexte et les enjeux du projet CAPTHOM sont décrits. Nous faisons ensuite le point sur la notion de détection de présence humaine. Les différentes technologies utilisées pour la détection de présence sont présentées en soulignant leurs avantages et leurs inconvénients. Ceci, dans l'objectif de retenir les technologies répondant au mieux aux exigences du dispositif CAPTHOM.
- Dans le second chapitre, la localisation 3D d'une personne à partir d'informations visuelles est étudiée. Nous faisons un résumé des méthodes existantes puis nous proposons une solution alternative basée sur l'estimation à horizon glissant. Le problème d'estimation de position 3D est reformulé en un problème d'optimisation non linéaire dans le plan image. Des résultats expérimentaux illustrent la faisabilité de l'approche.
- Le dernier chapitre est consacré à la fusion de données multi-capteurs pour des applications d'habitat intelligent. Nous montrons l'intérêt et la nécessité de fusionner

ces données aussi bien pour renseigner les divers registres (détection de présence, localisation 3D, détermination de la posture, nombre de personnes) que pour prendre les décisions finales (alerte en cas de surveillance de personnes âgées, déclenchement de l'éclairage ou du chauffage). Nous proposons une architecture de fusion pour le dispositif CAPTHOM avec différents niveaux de fusion. Nous détaillons le module de fusion de données pour la détermination de posture de la personne basée uniquement sur des informations extraites de l'image et donc sans capteur embarqué sur la personne.

Des résultats expérimentaux illustrent l'efficacité de l'approche développée.



# Chapitre 1

## La détection de présence humaine

Ce premier chapitre est dédié à la problématique de détection de présence humaine. Être capable de savoir si quelqu'un est dans une pièce ou non, est une information primordiale pour de nombreuses applications d'habitat intelligent. La détection de présence, trop souvent assimilée à tort à la détection de mouvement, est une notion de plus en plus utilisée pour des applications allant de la surveillance de bâtiments ou de personnes à la gestion du chauffage ou de la lumière.

Ce travail de thèse s'inscrit dans un projet du pôle de compétitivité  $S^2E^{21}$ , le projet CAPTHOM, qui a pour objectif de développer des capteurs intelligents de présence humaine.

Dans une première partie, nous situons ce travail de thèse dans son contexte industriel et précisons les enjeux du projet pour des applications de maintien à domicile ou de gestion énergétique d'habitat intelligent.

Ensuite, nous étudions en détail la notion de présence humaine. Nous faisons un tour d'horizon des technologies existantes en soulignant leurs avantages et leurs limitations. Ceci permet alors de justifier les choix technologiques faits pour le projet CAPTHOM et de soulever les points clés.

---

<sup>1</sup> $S^2E^2$  : Sciences et Systèmes de l'Énergie Électrique

## 1.1 Le projet CAPTHOM

Le projet CAPTHOM, pour Capteur d'Humain, est un projet du pôle de compétitivité  $S^2E^2$ , Sciences et Systèmes de l'Énergie Électrique. L'objectif de ce projet est de développer des capteurs intelligents pour détecter (ou non) la présence d'une personne. Deux applications sont plus particulièrement ciblées : d'une part, la gestion optimisée de l'énergie dans l'habitat industriel ou particulier, et d'autre part, la surveillance à domicile de personnes âgées ou handicapées.

Par ses deux applications, le projet CAPTHOM est lié à deux autres projets de plus grande envergure. Il s'agit du projet SESAME (Sciences Électriques et Systèmes Appliqués au Management de l'Énergie) et du Groupement d'Intérêt Scientifique MADONAH (Maintien et Assistance Diurne Ou Nocturne pour personnes Âgées ou Handicapées).

- Le projet SESAME s'intéresse à la mise au point d'équipements, à la fois professionnels et grand public, aptes à améliorer et optimiser la consommation énergétique dans tout type d'habitat. Trois axes principaux constituent l'ossature du projet SESAME : l'efficacité énergétique, le confort de l'habitation et, la sécurité et la santé des personnes. L'objectif visé est de concevoir la maison de demain en axant la recherche sur l'interopérabilité, la gestion à distance, l'adaptation à tout type d'environnement ainsi que l'ajout de nouvelles fonctionnalités.
- Le GIS MADONAH développe une plateforme technologique d'habitat intelligent au sein d'un établissement pour personnes dépendantes. L'objectif est de proposer la mise en place (développements et tests) de dispositifs orientés maintien à domicile dans un environnement habitable par des personnes âgées ou handicapées.

### 1.1.1 Le contexte du projet

Un consortium de six entreprises, pour la majorité des PME de la région Centre, est à l'origine du projet CAPTHOM auquel se sont ajoutés un CRITT (Centre Régional pour l'Innovation et le Transfert de Technologies), un pôle de compétences et un laboratoire de recherche. Les différents partenaires industriels et institutionnels sont donc :

- AGILICOM : développement et fourniture de produits et services en communication industrielle ;
- LEGRAND : produits et systèmes pour installations électriques et réseaux d'informations ;

- SOREC : fabrication de cartes électroniques assemblées, porteur du projet CAPTHOM ;
  - ST MICROELECTRONICS : développement, fabrication et commercialisation de puces électroniques (semi-conducteurs) ;
  - THERMOR : conception de chauffage et chauffe-eau électriques ;
  - WIRECOM Technologies : conception et développement de solutions permettant de maîtriser les énergies dans les bâtiments et habitations ;
- 
- CRESITT Industrie : CRITT en électronique de la Région Centre ;
  - Le Pôle Capteurs Automatismes dont l'un des rôles est de relier des projets industriels innovants de la Région Centre et des laboratoires de recherche ;
  - L'institut PRISME (Institut Pluridisciplinaire de Recherche en Ingénierie des Systèmes, Mécanique et Énergétique) de l'Université d'Orléans.

Au regard des activités de ces différentes entreprises, on comprend aisément l'intérêt que représente la conception d'un dispositif intelligent de détection de présence humaine. Pour épauler les entreprises et proposer des solutions technologiques et scientifiques aux problématiques dans l'avancement du projet, l'Institut PRISME est présent au travers de 4 thèses et constitue la partie recherche du projet CAPTHOM. Afin de bien cerner la problématique soulevée par cette thèse et de dessiner les liens qui la lie aux autres, il nous semble intéressant de présenter succinctement les trois autres sujets de recherche.

Plusieurs questions ont été soulevées lors des réflexions préalables au lancement du projet CAPTHOM.

### **Comment peut-on modéliser la présence humaine et la caractériser par des signaux ?**

C'est l'objet d'une thèse qui a pour but de proposer une modélisation de la présence humaine en s'appuyant sur des technologies existantes (capteur pyroélectrique, détecteur ultrason, ...) et en envisageant des améliorations pour chacune de ces technologies [Bel08].

### **Comment déterminer automatiquement la présence d'une entité humaine dans une image ?**

Parmi le panel des technologies pouvant délivrer des informations relatives à la présence humaine, la vision est une des technologies que le projet CAPTHOM ne pouvait ignorer. Un travail de recherche a permis de mettre au point un algorithme de traitement d'images [Ben09a][Ben08]. Cet algorithme détecte une entité dans une image et la carac-



térise comme étant humaine ou non.

### **Comment optimiser le placement des différents capteurs afin d'assurer une couverture spatiale et temporelle ?**

Comme nous le verrons par la suite, l'information de la présence humaine ne peut être obtenue par un seul capteur. Il faut donc disposer de plusieurs mesures issues de différents capteurs. Se pose alors le problème du positionnement « optimal » de ces capteurs, abordé dans une troisième thèse [Dav09]. L'optimisation du placement d'un réseau de capteurs a été étudiée et a donné naissance à un outil de simulation nommé SNOOPS (Sensor Network : Optimisation Of Placement by Simulation) [Dav07]. La première fonctionnalité de SNOOPS est de pouvoir simuler les possibilités de placement des différents capteurs dans un environnement donné (que l'utilisateur peut lui-même configurer) et d'obtenir le meilleur placement envisageable selon que l'on cherche à optimiser la couverture de la zone surveillée ou bien à minimiser les coûts. La seconde fonctionnalité de SNOOPS réside dans la possibilité, une fois le placement des capteurs effectué, de simuler des scénarios décrivant le déplacement et le comportement d'une personne afin d'obtenir un aperçu des réponses de chacun des capteurs à chaque phase du scénario.

### **Comment utiliser l'ensemble des informations de natures différentes pour répondre aux attentes en matière de gestion énergétique et de maintien à domicile ?**

En considérant le placement de capteurs optimisé, la multiplicité des technologies, à la fois en termes de quantité et de variété, conduit à une multitude d'informations à traiter. Il apparaît nécessaire de développer une méthode de traitement et de combinaison de ces informations suivant une architecture intelligente et efficace : ceci pour aider à la prise de décision relative à la présence ou non d'une personne dans un environnement donné. C'est l'objet de la présente thèse qui s'articule, autour de la fusion de données provenant de capteurs multiples et variés, pour la détection de présence humaine, la localisation 3D et la détermination de la posture d'une personne.

La conception et la structure d'un dispositif de fusion dépendent non seulement des objectifs visés, objectifs que nous allons détailler au travers des enjeux de la détection de présence (§1.1.2) mais également des technologies retenues pour répondre à la problématique (§1.2).

## **1.1.2 Les enjeux de la détection de présence**

### **1.1.2.1 Les enjeux pour CAPTHOM**

Les différents partenaires industriels du projet CAPTHOM ont établi un cahier des charges exprimant les besoins et les exigences que le capteur intelligent doit satisfaire. Nous présentons ici les principales missions et objectifs qui ressortent du cahier des charges.

#### **1. Une détection de présence continue et fiable**

Le dispositif CAPTHOM doit être capable de détecter la présence d'un humain (et pas uniquement le mouvement) dans un environnement surveillé. Cette détection doit se faire à coup sûr tout en minimisant les fausses et les non-détections. La notion de détection de présence doit être comprise comme la possibilité d'affirmer qu'il y a occupation ou non de l'environnement considéré, et qu'en cas d'occupation, on est capable de discriminer un humain de toute autre entité.

#### **2. Une amélioration de l'existant**

Des technologies sont déjà utilisées dans les projets d'habitat intelligent voire développées par des partenaires du projet CAPTHOM (l'entreprise LEGRAND commercialise déjà des détecteurs infrarouges passifs pour la gestion de l'habitat). Ces technologies doivent donc servir de base au développement du dispositif CAPTHOM. Des améliorations des composants et matériaux existants ainsi que le développement de traitements plus efficaces de leurs données doivent être envisagés.

#### **3. Un dispositif non invasif et non intrusif**

Contrairement au panel de capteurs que l'on retrouve dans la plupart des projets similaires, aucun capteur ou détecteur ne doit être porté par la personne, cette dernière devant en effet rester un élément passif du dispositif (capteur non invasif). L'objectif est d'être le plus transparent pour la personne et de modifier le moins possible son habitat si l'on travaille sur de l'existant (capteur non intrusif). Par conséquent, le module final devra être aussi peu encombrant que le sont les détecteurs infrarouges passifs actuellement commercialisés, mais avec de meilleures performances.

#### **4. Une approche sous deux formes**

Deux concepts doivent émerger des réflexions et développements :

- Le premier est destiné à la simple détection de présence humaine pour la commande de matériels énergétiques. Il doit donc obéir à des contraintes en termes de coût et de temps de réponse afin de s'aligner sur les détecteurs infrarouges passifs, couramment

utilisés en gestion de l'habitat ;

- Le second concept, sous une forme plus évoluée, intégrera une composante vision et sera destiné à des applications de maintien à domicile (surveillance de personnes âgées avec détection des situations de chute).

## 5. Autres besoins

Enfin, outre la présence, un certain nombre d'informations doit être consultable à n'importe quel moment. Parmi ces informations, renseignées dans des registres, on retrouve la localisation (2D et 3D) de la personne, le nombre de personnes présentes dans l'environnement et la posture de la personne. Suivant l'application considérée, le dispositif utilisera tout ou une partie des registres renseignés.

Les enjeux du projet CAPTHOM sont donc de concevoir un dispositif de **détection de présence**, de **localisation**, de détermination de la **posture d'une personne** et du **nombre de personnes** présentes, et d'apporter ainsi des solutions pour deux types d'application : la **gestion de l'énergie** dans les bâtiments tertiaires et dans les résidences individuelles, et **l'aide au maintien à domicile**.

***Remarque 1.1** Grâce au GIS MADONAH, le projet CAPTHOM bénéficie d'une plateforme d'habitat intelligent de 100m<sup>2</sup> située dans un établissement pour personnes dépendantes, « Les résidences Bellevue » à Bourges. Cette plateforme comporte les pièces essentielles d'un habitat conventionnel et regroupe un ensemble d'équipements domotiques et de réseaux de capteurs destiné à collecter les données utiles à la reconnaissance de comportements défaillants et à la bonne gestion d'un dispositif d'alertes. L'intérêt d'une telle plateforme pour le projet CAPTHOM est de pouvoir tester, dans des conditions réelles, le comportement des capteurs retenus et de corriger ou affiner les algorithmes développés. Cependant, cette plateforme n'est actuellement pas équipée de caméras pour des raisons éthiques et n'a donc pas permis, pour le moment, une validation expérimentale de l'approche proposée.*

### 1.1.2.2 Le contexte national et international

Depuis ces dernières années, de nombreux projets d'habitat intelligent, proches des préoccupations de CAPTHOM, se sont développés sur le plan national et international. Que ce soit pour répondre à des préoccupations environnementales et de confort grâce à des dispositifs de gestions énergétiques [Kid99] ou pour améliorer l'autonomie et la surveillance de personnes âgées [Ric93][Nou00], il est nécessaire de posséder une information relative à la présence humaine. Nous présentons ci-dessous, et de manière non exhaustive,

quelques projets d'habitat intelligent proches de la problématique CAPTHOM.

L'habitat intelligent, (Smart Home en anglais), peut être défini comme étant une intégration complexe de technologies, de matériels électriques ou thermiques, de télécommunications, . . . Parmi les projets d'habitat intelligent, on distingue deux applications différentes : la gestion énergétique et l'aide au maintien à domicile de personnes âgées. La gestion optimisée de l'énergie consiste à fournir aux habitants tout le confort possible à l'aide de capteurs dissimulés dans l'habitat. On retrouve de nombreux projets notamment aux Etats-unis qui s'intéressent à cette problématique du confort et de la gestion de l'énergie sous la forme d'habitat communicant [Kid99][Mas99]. En Europe, on peut citer des projets comme le projet E2Home entre Ericson et Electrolux [Yeo08] ou encore le projet Homelab de Philips. Enfin, au niveau national, la société Legrand, au travers de sa maison domotique, propose ses nouveaux produits permettant de renouveler l'appareillage électrique, ses usages ainsi que la distribution des données [Leg05]. L'idée, avec l'utilisation des produits InOne by Legrand, est de gérer de manière personnalisée l'éclairage de la maison, le chauffage ou l'alimentation des appareils électroménagers en offrant la possibilité de programmer des scénarios suivant l'ambiance et les économies que l'on veut réaliser. Dans ces projets d'habitat communicant, deux types de données sont considérés : les données relatives à l'habitat et les données relatives à l'habitant. Si l'acquisition des données relatives à l'habitat est réalisable à partir de capteurs (température, luminosité), les données relatives à l'habitant sont plus difficiles à obtenir. Il s'agit en effet de surveiller son comportement, ses interactions avec son habitat et ses caractéristiques (par exemple, les caractéristiques physiologiques).

L'habitat intelligent pour la santé noté HIS, (Health Smart Home en anglais), s'est développé pour aider le maintien à domicile de personnes âgées ou handicapées. Dans ces projets, un grand nombre de capteurs différents interviennent pour observer les paramètres médicaux et physiologiques d'une part mais également les paramètres comportementaux et environnementaux d'autre part. Pour une approche plus approfondie des HIS, le lecteur intéressé trouvera des états de l'art plus complets dans [Ria01][Nou03][Cha08].

- Au sein du laboratoire TIMC-IMAG de Grenoble, de nombreuses recherches s'articulent autour de la conception d'HIS [Dem02] [Nou04][Fle08b][Vac09]. Un appartement a été conçu afin de mettre en oeuvre les méthodes développées autour du maintien à domicile. Une approche multi-capteurs disposés à la fois dans l'appartement test et sur la personne est considérée :

- des microphones, pour la reconnaissance de paroles ;

- un réseau de détecteurs infrarouges passifs pour la détection de mouvement ;
- un accéléromètre embarqué sur la personne délivrant des informations sur l'activité de la personne ;
- une station météorologique ;
- des capteurs physiologiques.

Les objectifs des recherches menées au sein de cet HIS, sont entre autres, de surveiller l'état de santé de la personne, d'améliorer le niveau d'autonomie, de remédier à la détérioration des facultés cognitives, de détecter les déplacements et les chutes éventuelles. Des réflexions sont également menées sur l'impact d'un dispositif d'habitat intelligent sur la personne et sur les limites éthiques à respecter.

- Le projet PROSAFE, développé par le LAAS-CNRS de Toulouse, a travaillé au développement d'un système de surveillance de personnes âgées pour répondre à deux objectifs : contribuer au maintien à domicile en assurant une surveillance constante et générer des alarmes en cas de situations anormales (longue période d'immobilité, agitation). Le système se compose d'un réseau de capteurs pyroélectriques reliés à un poste distant [Cha03][Bon08b].

- Dans le cadre du projet CompanionAble, le système de telemonitoring EMUTEM (Environnement MULTimodal pour la TELévigilance Médicale) a été développé [Med08] [Med09]. Ce système rassemble trois dispositifs : (1) ANASON pour la gestion des microphones [Ist06], (2) RFPAT pour la collecte des données provenant des capteurs physiologiques placés sur la personne [Bal04] et (3) GARDIEN qui gère le réseau de capteurs pyroélectriques déployés dans l'appartement test [Ban03]. L'objectif du système EMUTEM est de délivrer en sortie deux informations : une information « alarme » relative à la détection ou non d'une situation de danger ou d'alerte, et une information de « localisation » relative à l'occupation des pièces de l'appartement.

- D'autres projets de telemonitoring existent comme le projet TelePat qui analyse des données physiologiques et relatives à l'activité [Bou06], le projet Tandem, suite du projet Telepat, qui ajoute l'utilisation d'accéléromètres ou encore le projet DESDHIS (Détection de Situations de Détresse dans un Habitat Intelligent Santé) avec un dispositif de télémédecine s'appuyant sur l'utilisation de capteurs sonores, de capteurs de déambulation et d'activité [Fle08a].

En résumé, à travers l'étude de ces projets, on constate que certaines notions et attentes

reviennent systématiquement dans la conception d’habitat intelligent que ce soit pour la gestion énergétique de l’habitat ou le maintien à domicile. La détection de la présence de la personne constitue l’une des premières informations essentielles. Elle est généralement réalisée à l’aide de détecteurs de mouvement tel que des détecteurs infrarouges passifs. La détermination de l’activité de la personne, afin de détecter des situations d’agitations ou de chute, correspond à la seconde exigence pour les projets d’HIS. Les capteurs embarqués sur la personne, comme les accéléromètres ou les capteurs physiologiques, sont très souvent utilisés.

Au sein du projet CAPTHOM, l’objectif est d’assurer une détection de présence humaine à tout instant ainsi que la détermination de diverses informations comme la posture et la localisation sans avoir recours à des dispositifs embarqués sur la personne.

Nous allons maintenant préciser la notion de présence humaine, puis présenter les différentes technologies existantes pour la détection de présence humaine en soulignant leurs avantages et leurs inconvénients.

## **1.2 La détection de présence humaine**

Tout d’abord, il nous a semblé important de préciser le plus clairement possible la notion qui se cache derrière le terme détection de présence, trop souvent assimilée à de la détection de mouvement. Détecter la présence humaine, c’est être à la fois capable de détecter le mouvement mais également l’absence de mouvement, lorsque la personne est immobile. En effet, le dispositif de détection de présence humaine doit être opérationnel même si la personne ne réalise plus des mouvements suffisamment significatifs pour être détectée (travail à un bureau ou chute accidentelle).

Nous allons détailler, dans un premier temps, les points essentiels qui caractérisent la problématique de détection de présence humaine.

### **1.2.1 La problématique de détection de présence humaine**

Ces dernières années ont vu le développement de dispositifs de détection de présence au sein de projets plus globaux d’habitat intelligent [Nou00][Med09]. A une plus petite échelle, on a constaté également que de plus en plus de personnes cherchaient des solutions afin de sécuriser leur domicile et/ou d’automatiser au maximum la gestion de leur habitat. On a alors vu arriver sur le marché des dispositifs aptes à répondre à ces besoins, généralement décrits comme étant des systèmes de détection de présence. Cependant, en

examinant de plus près la technologie utilisée dans ces systèmes, on remarque, dans la majorité des cas, que ces dispositifs de détection de présence ne sont en fait que des détecteurs de mouvement. En effet, les termes de « détection de présence » et de « détection de mouvement » sont souvent utilisés de manière équivalente alors que la présence humaine ne se limite pas aux mouvements que peut effectuer une personne mais est une notion beaucoup plus complexe.

Certes la capacité d'un système à détecter les mouvements et/ou les déplacements d'une personne est nécessaire si l'on travaille à la détermination de la présence d'un humain dans un environnement, mais elle n'est en aucun cas suffisante. Il arrive que l'on reste immobile ou quasi immobile plus ou moins longtemps au cours de la journée, notamment lorsque l'on est assis devant son bureau au travail ou chez soi. Si l'on étudie par exemple, un dispositif dit de « détection de présence » servant à gérer l'éclairage, on constate que tant que la personne effectue des mouvements suffisamment amples dans le champ de détection du capteur alors la minuterie gérant l'allumage de la lampe est relancée. A l'inverse, si la personne cesse de bouger ou si elle n'effectue plus de mouvements suffisamment prononcés, le détecteur considère qu'il n'y a plus personne dans la zone surveillée, la lumière s'éteint contraignant la personne à bouger dans la direction des faisceaux de détection du capteur pour signaler sa présence. On constate, à travers ce court exemple, qu'un dispositif de détection de présence humaine doit également être capable de déceler « l'absence de mouvement » et de l'interpréter.

Au problème de non détection que nous venons de soulever peuvent s'ajouter des problèmes de fausses détections si l'on se limite à de la détection de mouvement. En effet, dans la notion de détection de présence humaine, on retrouve le terme « humain », ce qui implique que le dispositif de détection ne doit envoyer un ordre que s'il détecte une cible humaine. Or, avec un système de détection de mouvement, il peut arriver qu'une entité autre qu'une entité humaine (un animal par exemple), passe dans le champ de détection du capteur et déclenche ainsi la mise en marche du chauffage. On se retrouve alors avec un dispositif qui, au lieu d'améliorer la gestion énergétique d'une habitation, n'entraîne aucune différence en terme de consommation voire dégrade la situation antérieure.

La détection de personnes en mouvement ou immobiles et l'identification d'une entité humaine constituent les principales propriétés que l'on doit retrouver dans un dispositif de détection de présence humaine. A l'heure actuelle, des technologies offrent déjà des solutions en réponse à certaines de ces propriétés recherchées mais aucune, à notre connaissance, prise de manière individuelle, ne permet de répondre à l'ensemble des be-

soins.

## 1.2.2 Les solutions technologiques et leurs limitations

Les technologies basées sur le rayonnement infrarouge [Ria00] d'une part et sur la capture et le traitement d'images [Sch09] d'autre part constituent les principales approches rencontrées dans les projets d'habitat intelligent. Ces technologies permettent d'apporter des solutions pour certains des problèmes soulevés précédemment. D'autres approches sont parfois envisagées soit de manière indépendante, soit combinées avec une des deux technologies précédemment citées. Toutes possèdent leurs avantages et leurs inconvénients comme on va le voir ci-dessous.

### 1.2.2.1 Les technologies basées sur le rayonnement infrarouge

Des études portant sur les rayonnements électromagnétiques ont montré que tout corps porté à une température supérieure au zéro absolu ( $0^{\circ}\text{K}$ , soit  $-273,15^{\circ}\text{C}$ ) émet un rayonnement électromagnétique que l'on appelle rayonnement thermique ou encore rayonnement du corps noir. Un corps noir est, par définition, un objet idéal qui aurait pour propriété d'absorber toute l'énergie électromagnétique qu'il recevrait, sans en réfléchir ni en transmettre. Si on augmente sa température, un tel objet finira par émettre de la lumière. Le rayonnement émis par un corps noir dépend de sa température et peut ainsi se situer dans toutes les longueurs d'onde. Deux lois permettent de définir et caractériser le rayonnement d'un corps noir. La première est la loi de Wien caractérisant la dépendance du rayonnement d'un corps noir à la longueur d'onde. La seconde est la loi de Planck permettant de définir la distribution de la luminance énergétique monochromatique du rayonnement thermique du corps noir en fonction de la température. Il est à noter qu'un objet se comporte rarement comme un corps noir dû au fait qu'il réfléchit une partie de l'énergie électromagnétique et en transmet une autre partie mais n'absorbe jamais la totalité.

On peut assimiler le corps humain à un corps noir de température moyenne  $37^{\circ}\text{C}$  (soit environ  $310^{\circ}\text{K}$ ). L'énergie rayonnée d'un corps humain en fonction de la longueur d'onde est donnée par la loi de Planck et se situe dans l'infrarouge : la plus grande partie est émise entre  $6\mu\text{m}$  et  $14\mu\text{m}$ . La loi du déplacement de Wien stipule que la longueur d'onde à laquelle un corps noir émet le plus de flux lumineux énergétique est inversement proportionnelle à sa température en Kelvin :

$$\lambda_{max} \cong \frac{3 \cdot 10^{-3}}{T}. \quad (1.1)$$



Pour une température de  $310^{\circ}\text{K}$ , on obtient alors une longueur d'onde maximale de  $9,7\mu\text{m}$ . Le rayonnement émis par le corps humain est donc un rayonnement infrarouge (noté rayonnement IR dans la suite du manuscrit).

On distingue trois technologies basées sur l'utilisation du rayonnement thermique :

- les capteurs pyroélectriques ;
- les détecteurs infrarouges passifs ;
- les détecteurs thermoélectriques.

### Les capteurs pyroélectriques

Les capteurs pyroélectriques (figure 1.1) mesurent le rayonnement thermique produit par une entité vivante. Ces capteurs sont sensibles à des brusques variations du flux thermique incident, et uniquement à des variations. En présence de flux constant, le capteur pyroélectrique ne réagit pas. Cette limitation est due au fait que le capteur pyroélectrique ne réagit que lors de l'apparition d'une modification par rapport à un état de référence, mis à jour de manière permanente. Ceci peut être source de problèmes pour une application de détection de présence.



FIG. 1.1 – Représentation d'un capteur pyroélectrique.

En effet, si une personne reste immobile devant le capteur pendant un certain temps, ce dernier finit par considérer que la personne fait partie intégrante de l'environnement de référence et par conséquent il ne la détecte plus. Tant que la personne restera immobile, elle sera invisible pour le capteur pyroélectrique, situation qui peut s'avérer gênante suivant l'utilisation que l'on en fait. Ce n'est qu'au moment où la personne va bouger de nouveau que le capteur va la détecter et envoyer de nouveau un signal.

D'autre part, l'homme n'est pas la seule source de flux thermique dans une habitation et n'est donc pas toujours le seul responsable du déclenchement d'un capteur pyroélectrique. Il peut arriver en effet que l'entrée d'un animal de taille moyenne, comme un chien, entraîne l'activation du capteur. Une solution envisageable pour contourner ce problème est de positionner les capteurs de telle manière qu'ils ne détectent pas de petites entités

ou en les disposant à différentes hauteurs.

Les sources de chaleur comme les radiateurs ou le rayonnement solaire sur une surface vitrée peuvent venir perturber le comportement du capteur, ce dernier fournissant alors des informations erronées.

***Remarque 1.2*** *Le capteur pyroélectrique est un détecteur de mouvement mais ne peut pas être utilisé seul, pour la détection de présence, qui plus est de présence humaine.*

### **Les détecteurs infrarouges passifs**

Les détecteurs infrarouges passifs (détecteur IRP) combinent des capteurs pyroélectriques et des lentilles de Fresnel. Ces dernières permettent d'orienter le champ de vision et de définir des cônes de détections spécifiques suivant la lentille utilisée. Les lentilles de Fresnel découpent l'environnement surveillé par le capteur, en zones de détection et en zones de non détection. On parle alors de lobes de détection qui sont caractérisés par une direction, une largeur angulaire et une portée (figure 1.2). Les performances des détecteurs IRP dépendent du nombre, de la forme et de la répartition des lobes de détection. Si la personne se trouve dans une des zones de détection, les lentilles de Fresnel focalisent une partie du flux IR sur la surface du capteur pyroélectrique. À l'inverse, si la personne se trouve dans une zone aveugle, aucune information n'est reçue par le capteur. Lors d'un déplacement, une personne parcourt alternativement des zones de détection et des zones de non détection, créant alors des variations du flux IR incident et engendrant un signal.

Dans un état de l'art des technologies opérant dans l'infrarouge [Esc05], les détecteurs IRP sont présentés comme des détecteurs de mouvement efficaces et sont largement commercialisés (figure 1.3). Plusieurs variantes leur permettent de répondre à certaines difficultés ou d'apporter un complément d'information.

- Associés en réseau, ils permettent d'effectuer une localisation grossière de la personne dans l'habitat [Ban03].
- Des détecteurs IRP spécifiques ont aussi été conçus avec des limitations dans le domaine spectral. Seul l'ensemble des longueurs d'ondes correspondant au rayonnement d'un humain est utilisé grâce à un filtre optique. De plus, dans un souci de diminuer les fausses détections, ces capteurs dédiés à la détection de mouvement sont généralement des capteurs doubles, c'est à dire qu'ils se composent de deux éléments sensibles avec des polarités inversées [Gob06].

Comme le capteur pyroélectrique, le détecteur IRP n'assure pas la détection en cas de flux constant et est sensible à des flux thermiques perturbateurs.

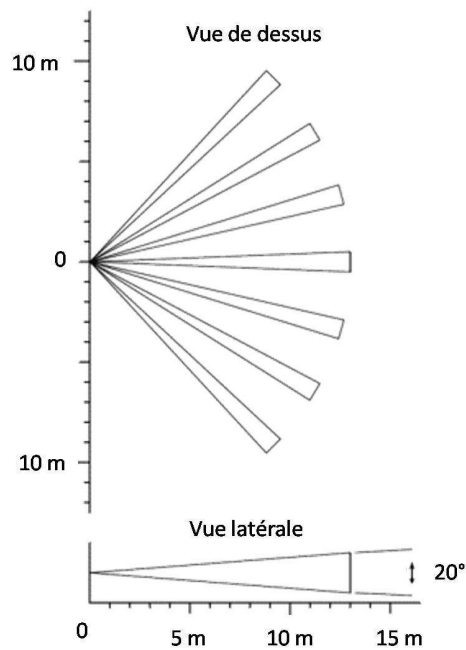


FIG. 1.2 – Lobes de détection d'un détecteur infrarouge passif.



FIG. 1.3 – Exemple de détecteurs infrarouges passifs commercialisés.

**Remarque 1.3** *Les détecteurs IRP sont sensibles à la façon dont la personne est habillée. Un exemple simple est l'entrée d'une personne portant un manteau en hiver. La température extérieure étant généralement très inférieure à la température ambiante, le manteau couvrant la personne constitue un corps froid parasitant le rayonnement thermique de la personne. En conséquent, cette dernière peut ne pas être détectée.*

### Les détecteurs thermoélectriques

Les détecteurs à effet thermoélectrique, comme les thermopiles, semblent constituer une des pistes de recherche les plus étudiées concernant le rayonnement infrarouge. Il nous

semble intéressant de décrire brièvement le mécanisme de cette technologie.

La technologie thermoélectrique exploite des phénomènes de polarisation spontanée sous l'action d'une différence de température imposée aux extrémités d'un matériau. Cette technologie repose sur l'effet Seebeck, qui montre qu'une différence de potentiel apparaît à la jonction de deux matériaux soumis à une différence de température (figure 1.4).

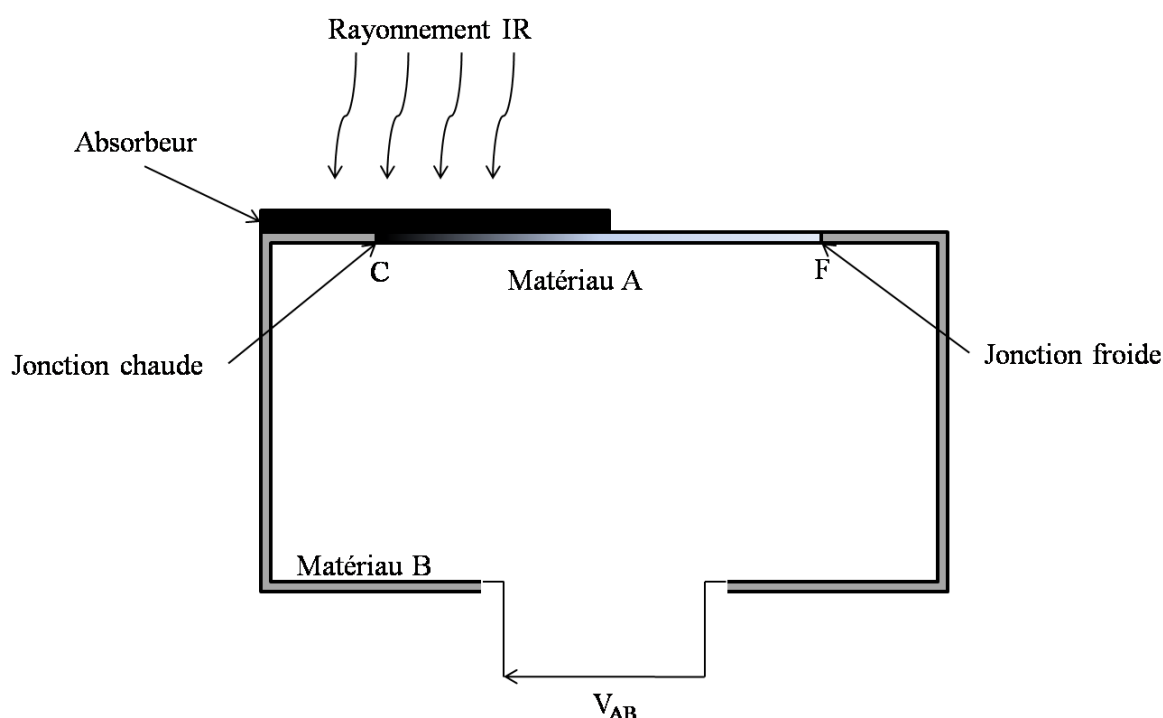


FIG. 1.4 – Schéma de l'effet Seebeck.

On dispose de deux matériaux  $A$  et  $B$  de nature différente, reliés par deux jonctions (une jonction chaude  $C$  et une jonction froide  $F$ ). Ces deux segments joints forment alors un circuit. Si l'on applique une différence de température  $\Delta T$  entre les jonctions  $C$  et  $F$ , une force électromotrice  $V_{AB}$  apparaît. Dans la pratique, une température de référence  $T_F$  est maintenue au niveau de la jonction froide tandis que la température  $T_C$  au niveau de la jonction chaude évolue en fonction des variations du flux IR incident auxquelles elle est soumise. Le gradient thermique  $\Delta T = T_C - T_F$  engendre une force électromotrice  $V_{AB} = \alpha_{AB} \cdot \Delta T$ , qui est fonction de l'écart thermique entre les deux jonctions mais également des propriétés électrothermiques  $\alpha_{AB}$  du couple de matériaux, que l'on appelle également thermocouple.

Afin d'augmenter la tension générée, plusieurs thermocouples sont généralement connectés en série. Cet ensemble de thermocouples constitue la thermopile. Pour une structure de  $N$  thermocouples, on obtient alors l'équation suivante :

$$V_{AB} = N.\alpha_{AB}.\Delta T. \quad (1.2)$$

**Remarque 1.4** *Les thermopiles sont utilisées à l'heure actuelle pour des mesures de température de proximité. On retrouve par exemple cette technologie dans des thermomètres auriculaires ou encore dans des capteurs de température utilisés dans le domaine aéronautique ou automobile.*

Les thermopiles, contrairement aux détecteurs IRP classiques, permettent de détecter la présence d'une personne immobile d'où l'intérêt de cette technologie. Des travaux de recherche ont été initiés dans ce domaine [Esc05].

L'utilisation des thermopiles pour la détection de présence amène l'apparition de nouvelles difficultés :

- comme pour les détecteurs infrarouges passifs, les thermopiles sont sensibles aux variations de la température ambiante qui peuvent venir perturber la température de la jonction froide, rendant difficile la définition d'une tension de référence au repos ;
- le signal délivré par le capteur est très faible et nécessite d'être amplifié, ajoutant alors des perturbations au signal d'origine ;
- la thermopile rencontre des problèmes d'atténuation du signal en fonction de la distance qui la sépare de la cible.

#### 1.2.2.2 La caméra vidéo

Affirmer, à partir d'une image, qu'une personne est présente ou non, s'avère relativement simple à partir du moment où l'on parvient à identifier et reconnaître une personne. L'image se révèle être une source d'informations très riche pour la détection de présence humaine. S'il est facile pour un humain de reconnaître un autre humain sur une image, la tâche est nettement moins triviale lorsque l'on développe un système automatisé. C'est pourquoi la reconnaissance et la caractérisation d'une entité comme étant humaine à partir d'informations visuelles ont fait et font encore aujourd'hui, l'objet de nombreuses recherches. N'étant pas des spécialistes du traitement d'images, notre objectif n'est pas d'établir un état de l'art exhaustif des méthodes de vision pour la détection humaine.

Le lecteur intéressé trouvera un état de l'art plus détaillé dans [Sch09]. Néanmoins, nous présentons brièvement les approches majeures en traitement d'images pour la détection de personnes afin de souligner les difficultés et/ou inconvénients. Elles seront un peu plus détaillées dans le Chapitre 2.

La méthodologie de détection de personne est sensiblement la même, que l'on travaille avec une image ou avec une vidéo (une vidéo n'étant en fait qu'une succession d'images auxquelles on applique le même procédé). Toutefois, le traitement d'une vidéo est un peu moins contraignant et moins gourmand que le traitement d'une image seule. Si l'on ne dispose que d'une image, le dispositif de détection n'a aucune connaissance a priori et doit donc parcourir l'intégralité de l'image à l'aide d'une fenêtre de détection. Pour une vidéo en revanche, le mouvement, schématisé par une soustraction d'arrière plan, différence entre l'image courante et l'image de référence représentant l'environnement sans humain, permet de restreindre la recherche dans les zones où un changement s'est produit. Cette restriction du champ de recherche diminue de ce fait le temps de calcul. Un mouvement suit généralement une trajectoire : il est alors possible d'utiliser les positions antérieures de la personne dans l'image pour prédire sa position future et ainsi anticiper la zone de recherche future. Le traitement appliqué pour la phase de caractérisation est ensuite le même que l'on travaille avec une image ou une vidéo.

On distingue, dans la littérature du traitement d'images, deux approches pour caractériser une cible humaine. La première approche regroupe les méthodes basées sur des modèles 2D ou 3D représentant, soit certaines parties d'une personne, soit l'intégralité de la structure du corps humain [See06][Wu 07]. La seconde approche regroupe les méthodes basées sur l'utilisation d'une fenêtre glissante et de techniques d'apprentissage supervisé [Pap00][Bel00].

- Les méthodes basées sur les modèles ont été utilisées pour diverses applications : la détection de personnes debout dans une scène de trafic [Lei05], la détermination de personnes dans des vidéos sportives [Ram06] ou encore dans [And08] où une personne est représentée par un modèle basé sur les membres et les articulations.
- Pour les méthodes basées sur un apprentissage supervisé, l'idée est de balayer l'image avec des fenêtres de tailles différentes. Deux critères doivent être pris en compte : la composante *aspect* qui code l'apparence visuelle de la personne et le *classifieur* qui détermine indépendamment pour chaque fenêtre si elle contient ou non un humain. Les méthodes basées sur les ondelettes de Haar, proposées la première fois par [Pap98] et que l'on retrouve dans [Vio05] et [Ben09b] peuvent être utilisées.

L'utilisation d'histogrammes de gradients orientés a également donné des résultats satisfaisants [Dal05].

La détection automatisée d'une personne dans une image pose des difficultés respectivement liées à la personne et à la qualité de l'image.

La personne constitue la première source de difficultés. En effet, le corps humain est une structure fortement articulée, qui peut prendre différentes configurations plus ou moins reconnaissables pour un système automatisé. Au cours d'un déplacement, la silhouette d'une personne change en permanence. Les classifieurs sont déterminés à partir de bases de données regroupant des images d'humains dans différents environnements et dans différentes positions. Il est difficile de balayer l'ensemble de l'éventail des positions : d'une part, car il est laborieux de constituer une base de données suffisamment complète, et d'autre part, parce qu'une trop grande diversité des positions peut entraîner une baisse de performance. Il est souvent judicieux de travailler avec des classifieurs spécifiques à certaines parties du corps ou liés à une posture particulière. De plus, les comportements et les caractéristiques diffèrent souvent d'une personne à l'autre. Les humains ne se déplacent pas de la même manière ; chacun ayant une vitesse et des mouvements qui lui sont propres. Chaque humain étant unique, ses caractéristiques (couleur de la peau, taille et corpulence) vont varier d'une personne à l'autre augmentant la difficulté de reconnaissance.

Concernant les difficultés liées au traitement d'images, on peut noter les occultations dues à des obstacles ou des vêtements qui dissimulent en partie ou en totalité une personne. Elles accentuent fortement la complexité du problème et constituent un des problèmes majeurs et incontournables en détection de personnes dans une image. A cela s'ajoutent les conditions d'éclairage auxquelles la qualité des images est très sensible.

***Remarque 1.5*** *On peut extraire de l'image des informations pertinentes pour la détection de présence, la localisation, le nombre de personnes présentes ou encore la posture de la personne. Se pose alors le problème du respect de la vie privée de la personne. L'image doit être considérée comme un signal (2D), traité et analysé. En aucun cas, l'image n'est visualisée sur un écran.*

### **1.2.2.3 Les autres technologies servant à la détection de présence**

Sont évoquées, dans cette partie, les autres technologies utilisées en tant que telles ou associées à d'autres pour la détection de présence.

## La détection volumétrique

La présence d'une personne entraîne des variations en matière d'occupation du volume d'une pièce. L'utilisation de techniques basées sur les ultrasons couplés à un dispositif de balayage peut être envisagée. Lorsque la pièce est vide, i.e. non occupée par un humain, des balayages réguliers permettent de construire une représentation de l'environnement surveillé en définissant les distances des différents éléments de la pièce par rapport au capteur. On établit alors une cartographie de référence qui donne une caractérisation de la scène au repos. Un individu pénétrant dans la pièce modifie le temps de parcours aller-retour des ondes envoyées dans sa direction et crée une nouvelle cartographie de la pièce. Les différences entre ces deux cartographies permettent de déterminer la présence ou non d'une entité dans la pièce.

Sur le même principe de fonctionnement, on peut également citer un autre capteur volumétrique : le capteur hyperfréquence, qui offre une portée et un volume de détection plus important. Ce dernier est souvent couplé à un détecteur infrarouge passif afin de minimiser les fausses détections.

***Remarque 1.6*** Comme pour les détecteurs IRP, l'utilisation de détecteur volumétrique ne permet pas l'identification de l'entité en mouvement comme étant une personne. Les détecteurs ultrasons peuvent être également sujets à des déclenchements intempestifs si ils sont perturbés par des ondes extérieures. Enfin, les ultrasons comme les capteurs hyperfréquence rencontrent des difficultés de détection pour des mouvements transversaux à leur axe d'orientation.

## La détection sonore

Les bruits ou les sons peuvent être révélateurs de la présence humaine dans un environnement. La mise en place de microphones dans un habitat permet de savoir si une personne est présente dans telle ou telle pièce mais également de déterminer si elle est dans une situation de danger. L'information des microphones nécessite ensuite un post-traitement. D'une part, il faut être capable de discerner les bruits produits par l'homme (déplacement d'objets et parole) des autres bruits présents dans une habitation ou provenant de l'environnement extérieur [Ist06]. D'autre part, dans une optique de détermination de situation de danger, il est nécessaire de reconnaître des plaintes ou des appels à l'aide par rapport à une conversation classique.

***Remarque 1.7*** Comme pour l'image, les enregistrements sonores et vocaux peuvent être perçus comme une atteinte à la vie privée.



## La détection par capteur de pression

Autre élément caractéristique d'une personne, le poids peut être utilisé au travers de capteurs de pression sous la forme de tapis ou de dalles recouvrant le sol d'une pièce. Ce type de système permet de connaître facilement la position de la personne dans une pièce mais se révèle très onéreux. De ce fait, cette technologie est rarement retenue dans les projets d'habitat intelligent. En maintien à domicile, il est parfois envisageable d'équiper des surfaces beaucoup moins vastes mais plus stratégiques telles que des toilettes, une salle de bain ou la zone autour du lit. L'objectif, dans ce cas, est de détecter une éventuelle chute de la personne dans des zones accidentogènes.

***Remarque 1.8** Cette technologie peut être envisagée dans de futurs habitats intelligents. Elle ne peut pas être utilisée dans un habitat existant sans d'importants travaux, ce qui ne correspond pas aux exigences du projet CAPTHOM.*

## Les capteurs embarqués sur la personne

Enfin, pour terminer cette liste non-exhaustive, et bien que cette technologie ne satisfait pas le cahier de charges du projet, il nous semble important de mentionner les différents capteurs pouvant être portés par la personne. Ils se présentent sous la forme de bracelet, de ceinture ou de patch. Des puces RFID (Radio Frequency IDentification) sont mises en place au sein d'un bracelet, permettant de détecter la présence et d'identifier la personne [Cas09]. Des accéléromètres et/ou des gyromètres sont utilisés afin de détecter des chutes et des activités inhabituelles [Bal04][Li 09][Bar05].

L'ensemble des technologies évoquées dans cette partie permettent toutes d'apporter des informations quant à la présence ou non d'une personne dans un environnement surveillé. Toutes possèdent également des limitations que ce soit en termes de performance, de condition d'utilisation, de coût d'installation ou de maintenance. Mais, il ne faut pas oublier que, pour toute problématique de détection de présence humaine, l'homme est au cœur du processus.

### 1.2.3 L'homme, centre du processus de détection de présence

Il est essentiel, dès la phase de spécification et de conception d'un projet d'habitat intelligent, de se poser des questions en matière de tolérance et d'acceptabilité par l'homme des technologies utilisées.

Que ce soit pour des applications domotiques ou de surveillance de personnes âgées ou handicapées, l'utilisation de tel ou tel dispositif ne doit pas venir perturber les habitudes de vie, ni menacer l'intimité et la vie privée des personnes. Des questions éthiques sont inévitablement soulevées dès que l'on implique l'homme notamment :

1. par l'observation des personnes à travers l'œil des capteurs ;
2. par la transmission des données à l'extérieur du domicile ;
3. par le traitement de ces données par des personnes extérieures.

Une technologie, aussi efficace soit-elle, peut être perçue chez certaines personnes comme une menace et une atteinte à leur vie privée. Cette prise en compte de l'acceptabilité et des préoccupations de l'homme face aux exigences techniques liées à un équipement de surveillance se retrouvent dans tous les projets qui traitent du maintien à domicile [Nou04] [Bon08a] [Le 08].

Dans la plupart des projets, on retrouve des détecteurs IRP en raison de leur faible impact sur la vie privée et leur faible encombrement (ils ne modifient que très peu l'environnement existant). Les capteurs embarqués sur la personne (capteurs invasifs) font leur apparition dans plusieurs projets HIS. L'usage de dispositifs vidéo ou sonore est encore difficile à faire accepter même si ils peuvent offrir de meilleures performances. Un subtil compromis entre efficacité et liberté doit être trouvé pour répondre aux attentes aussi bien des particuliers que des industriels.

## 1.3 Conclusion

Nous venons de voir au cours de ce chapitre que la détection de présence humaine constitue une problématique largement étudiée et joue un rôle prépondérant dans les applications d'habitat intelligent.

La détection de présence humaine est un phénomène complexe, impliquant la détection non seulement des mouvements mais également de l'absence de mouvement.

Un tour d'horizon des technologies utilisées dans les projets d'habitat intelligent a été présenté. Les forces et les faiblesses des technologies les plus couramment utilisées pour la détection de mouvement sont résumées dans le tableau 1.1.

A partir de ce tableau, nous avons retenu trois types de capteurs pour le projet CAP-THOM :

- les détecteurs IRP (pour la détection de mouvement) ;
- les thermopiles (pour la détection de présence) ;
- la caméra vidéo (pour l'identification, et ses nombreuses autres possibilités).

En effet, aucune technologie ne peut, seule, prendre en charge la détermination de la présence humaine de manière fiable. Il est donc nécessaire de concevoir un dispositif multi-capteurs impliquant de ce fait le développement d'une architecture de fusion de données. Cette dernière permet de traiter l'importante quantité d'informations disponibles et d'aider à la prise de décision en fonction de l'application. La fusion de données est abordée dans le troisième chapitre.

Concernant la caméra vidéo, elle est la technologie offrant le plus de possibilités pour renseigner les différents registres d'informations (détection, localisation, posture, nombre de personnes), soit directement par traitement d'images, soit indirectement par estimation ou fusion. La localisation 3D d'une personne, à partir d'une image, fait l'objet du chapitre suivant. Dans tous les cas, l'utilisation de la caméra se fait en respectant l'intimité de la personne. Ainsi, aucun enregistrement n'est effectué et aucune image n'est conservée ou visualisée. L'image est un support de l'information interne au dispositif CAPTHOM.

Types de détecteurs	Informations							Coût	Remarques
	Détection de mouvement	Détection de personnes immobiles	Nombre d'occupants	Localisation (2D ou 3D)	Activité	Posture	Discrimination		
Infrarouge passif	+	-	-	+/-	-	-	-	Faible	Localisation possible avec plusieurs capteurs.
Barrière(IR, US, lumineuse)	+	+/-	+/-	-	-	-	-	Faible	Possibilité de comptage entrée/sortie.
Ultrason	+	+	-	+/-	-	-	-	Faible	Difficile de connaître l'origine des détections.
Thermopile	+	+	-	+/-	-	-	-	Faible	Portée faible. Besoin d'amplification du signal.
Tapis détecteur de pression	+	+	+	+	-	-	+/-	Élevé	Installation coûteuse et intrusive.
Accéléromètre Gyromètre	+	-	-	-	+	+/-	-	Moyen	Technologie invasive.
Caméra vidéo	+	+	+	+	+	+	+	Faible à moyen	Nombreux algorithmes développés. Sensible aux variations lumineuses. Technologie intrusive.
Caméra infrarouge	+	+	+	+	+	+	+	Élevé	Plus de contraintes de luminosité mais coût prohibitif.

TAB. 1.1 – Tableau comparatif des détecteurs de présence humaine.



## Chapitre 2

# La localisation 3D par estimation visuelle sur horizon glissant

Afin de répondre aux exigences du cahier des charges du projet CAPTHOM, il faut pouvoir renseigner des registres d'informations parmi lesquels on trouve la localisation 2D et 3D de la personne.

La localisation 2D d'une personne à partir d'une image peut, de prime abord, paraître relativement aisée. C'est sans compter sur les problèmes d'occultations (présence d'obstacles) et les perturbations lumineuses dans l'environnement (éclairage, ensoleillement, reflet, ...).

La localisation 3D d'une personne, à partir d'une image, requiert en plus une information 3D de l'environnement.

Développées pour des applications de détection et de suivi de personnes dans l'image, des techniques basées sur l'utilisation d'une caméra stéréo ou de plusieurs caméras, ou sur la connaissance d'une transformation homographique apportent des éléments de réponse pour la localisation 3D. Compte tenu du coût du dispositif CAPTHOM à respecter, seule la transformation homographique, associée très souvent à un filtre de Kalman, a retenu notre attention.

En revanche, pour pallier certains inconvénients du filtre de Kalman étendu (hypothèse restrictive sur les bruits, linéarisation autour d'un point de fonctionnement, non prise en compte de contraintes), nous avons développé un estimateur visuel à horizon glissant. L'estimation de la position 3D est formulée en un problème d'optimisation non linéaire dans l'image sur un horizon passé glissant.

Dans une première partie, nous présentons brièvement les approches de localisation 3D existantes. Nous détaillons celle basée sur le filtre de Kalman étendu afin de montrer les éventuelles difficultés de mise en oeuvre et parce que notre approche y sera comparée dans un exemple.

Nous rappelons ensuite le principe de l'estimation d'état sur horizon glissant dans le cadre général non linéaire et nous présentons son extension à la localisation 3D d'une personne. L'élément clé de l'approche est le modèle global combinant un modèle de la caméra et un modèle non holonome de locomotion humaine.

Enfin, des résultats expérimentaux illustrent les possibilités de l'approche développée.

## **2.1 L'estimation de position en traitement d'images**

Le problème de localisation est étroitement lié au problème de détection et de suivi de cible, humaine ou non, dans un environnement 3D à partir de données visuelles 2D. Si la détection et la localisation d'une personne dans une image peuvent paraître triviales pour un humain, ces problématiques le sont beaucoup moins dès lors que l'on cherche à automatiser le traitement par un dispositif autonome. De nombreux travaux s'y consacrent depuis ces dernières années et nous présentons brièvement les principales approches.

### **2.1.1 Les approches de localisation 3D en traitement d'images**

Un dispositif de localisation 3D doit être capable de satisfaire certaines exigences parmi lesquelles on peut citer :

- une localisation dans des environnements complexes où des occultations peuvent survenir. Trois types d'occultation sont identifiables : (1) les occultations partielles où seule une partie de la personne est visible, (2) les occultations de courte durée où la personne n'est pas visible pendant un court moment, et (3) les occultations étendues où la personne quitte le champ de vision du capteur pendant une longue période ;
- une robustesse face aux problèmes de changement de luminosité, d'ombres, de réflexion dans un miroir et d'arrière plan complexe ;
- une reconstruction de la trajectoire par rapport à l'environnement 3D ou la surface du sol ;
- un temps de traitement de l'information rapide ;
- une facilité d'installation et de configuration.

Un dispositif capable de satisfaire l'ensemble de ces exigences n'existe malheureusement pas à notre connaissance. On retrouve dans la littérature du traitement d'image [Bro86][Phi91] et de la robotique mobile [Lee90][Wan92], des propositions de solutions qui permettent de répondre à certaines exigences.

**Remarque 2.1** *Les occultations et la robustesse du dispositif face aux diverses perturbations constituent les deux obstacles récurrents de ces méthodes.*

Le suivi de personnes (people tracking, en anglais) combine la localisation de la personne dans l'image et le suivi dans le référentiel monde, deux informations de dimensions différentes et obtenues par des phases différentes. Dans un premier temps, on détecte la présence d'une personne dans l'image. Puis, dans un second temps, on transmet les données à un dispositif de suivi qui estime la position dans l'environnement réel. La phase de suivi permet également d'effectuer une prédiction de la position au prochain instant d'échantillonnage. Cette prédiction permet alors de faciliter les calculs lors de la future phase de détection.

### **Phase 1 : la détection visuelle de personnes**

La problématique de la détection visuelle de personnes constitue un réel challenge en raison de la diversité des situations à prendre en compte et des environnements dans lesquels peut évoluer la personne. [Sch09] propose un état de l'art des méthodes les plus couramment utilisées en détection visuelle de personnes et les classe suivant deux catégories : (1) les méthodes basées sur des modèles de l'intégralité ou d'une partie du corps afin de déterminer la présence d'une personne et/ou son interaction avec l'environnement [Mik04][See06][Wu 07], (2) les méthodes basées sur un algorithme d'apprentissage permettant de déterminer des classifieurs et un balayage de l'image à l'aide de fenêtres de différentes tailles [Pap00][Bel00].

**Remarque 2.2** *Même si pour chacune de ces méthodes la détection de la personne est assurée de manière générale, certaines situations se révèlent encore problématiques : gestion des occultations ou encore des phénomènes d'ombre et de reflet. Le temps de calcul augmente très rapidement dès lors que l'on travaille avec un modèle complexe du corps humain ou avec un nombre important de classifieurs.*

**Remarque 2.3** *Selon [Sch09], l'amélioration des performances globales des approches de détection passe par une prise en compte du mouvement de la personne dans le processus de détection. Les deux problématiques, détection et suivi, sont encore trop souvent abordées*



de manière séparée même si certaines approches commencent à intégrer une phase de suivi dans la boucle de détection [Vio05][Dal06][Ben08].

## **Phase 2 : le suivi de personnes**

L'étude du suivi de personnes a fait l'objet de nombreuses recherches notamment pour la surveillance vidéo [And08] et en robotique mobile pour la navigation autonome dans un environnement complexe [Ess09][Baj09]. La première étape de toute phase de suivi consiste à extraire une information 3D de l'image afin de pouvoir réaliser l'estimation de position dans le repère 3D.

Une première approche consiste à utiliser un dispositif de stéréovision qui permet de délivrer une information de profondeur en plus de la position 2D dans l'image. L'utilisation de plusieurs caméras permet de générer des cartes de l'environnement dans lequel se trouve la personne et de gérer la majorité des problèmes d'occultations qui pourrait survenir. Dans [Baj09], les auteurs combinent les images provenant de deux caméras positionnées sur une voiture afin d'obtenir des mesures de profondeur qu'ils projettent sur une carte. La détermination de zones d'intérêt est réalisée par segmentation, puis une étape de classification, pour chaque zone, permet de savoir si un humain est présent ou non. Le suivi est réalisé par association de zones d'intérêt entre deux images successives.

**Remarque 2.4** *Le travail est réalisé avec des images de haute qualité ( $1024 \times 768$  pixels) et nécessite un dispositif de traitement performant pour s'assurer que l'application soit réalisable en temps réel. La projection de la position sur une surface 2D pendant l'étape de segmentation peut poser problème lorsque la personne est trop près d'un obstacle.*

Une seconde approche consiste à utiliser une unique stéréo caméra disposée en hauteur et observant la scène suivant un angle oblique. Dans [You08], les auteurs établissent également une carte, appelée *world-Z-map* à partir des données issues de la stéréo caméra. La *world-Z-map* correspond à l'enregistrement sur l'image de la distance de chaque point par rapport à la caméra. Les points les plus proches de la caméra ont les valeurs les plus importantes sur la *world-Z-map*. L'occultation, entre une personne et un obstacle ou entre deux personnes, est alors mieux gérée.

**Remarque 2.5** *La stéréo caméra tout comme le dispositif multi-caméras, sont des dispositifs coûteux et difficilement intégrables dans un dispositif embarqué. L'utilisation de telles approches n'est pas envisageable dans le projet CAPTHOM si l'on veut respecter les contraintes de coût, et nous impose donc l'utilisation d'une caméra perspective.*

La troisième approche consiste à utiliser une transformation homographique, qui est une projection bijective entre le plan image et le repère environnement, en considérant que la personne évolue sur une surface plane. Dans [Joh09], une homographie est utilisée pour déterminer la position de la personne dans l'environnement réel à partir de la position des pieds de la personne dans l'image. Lors de la phase de détection de présence, la personne est généralement englobée par une boîte : le milieu bas de la boîte englobante étant assimilé à la position des pieds. Un filtre de Kalman est ensuite utilisé pour réaliser l'estimation et la prédiction de la position dans l'environnement, c'est à dire le suivi de la personne [Rei79]. La présence de bruits sur les mesures visuelles, de déformations plus ou moins importantes selon la qualité de l'image et la prise en compte du modèle non linéaire du dispositif visuel justifient l'utilisation du filtre de Kalman étendu pour assurer le rejet de ces bruits et perturbations [Sat09][Bey99].

**Remarque 2.6** *Un avantage du suivi de personnes est de pouvoir gérer la présence de plusieurs personnes dans un même environnement. Un identifiant est attribué à chaque personne détectée. La gestion du suivi de plusieurs cibles nécessite une étape supplémentaire d'association des données avec la bonne cible. Notre application se limitant à du suivi d'une seule cible, le lecteur intéressé par le suivi multi-cibles (Multi-target tracking) trouvera un état de l'art dans [Sch03].*

### 2.1.2 L'estimation de position par filtrage de Kalman

Le filtre de Kalman étendu (Extended Kalman Filter – EKF), extension du filtre de Kalman développé dans les années 1960 [Kal60] à des systèmes non linéaires, est généralement utilisé pour estimer un vecteur d'état et/ou diminuer l'impact des perturbations et des bruits.

Dans le cas de l'estimation de la position 3D d'une personne, le vecteur d'état est la position  $(x, y)$  et l'orientation  $\theta$  de la personne. La sortie mesurée est le point d'intérêt dans l'image de coordonnées pixelliques  $(u, v)$ .

En adoptant le formalisme employé dans [Lip04], le modèle discret du processus s'écrit :

$$(M_d) \begin{cases} X_k = AX_{k-1} + w_k \\ \Omega_k = h(X_k) + v_k \end{cases} \quad (2.1)$$

où  $A$ , la matrice de transition, est une matrice bloc diagonale constante et  $w_k$  et  $v_k$  sont des bruits blancs gaussiens indépendants, de variance respective  $Q_k$  et  $R_k$ .  $\Omega_k = [u_k \ v_k]^T$  est le point d'intérêt dans l'image. Le vecteur d'état est  $X_k = [x_k \ \dot{x}_k \ y_k \ \dot{y}_k \ \theta_k \ \dot{\theta}_k]^T$ . La vitesse de déplacement est supposée constante sur un intervalle de temps  $T_e$  pour des

raisons de simplicité.

L'algorithme de l'EKF est un algorithme récursif qui se décompose en deux étapes.

a) une étape de prédiction qui fournit l'estimée a priori du vecteur d'état au prochain pas selon les équations suivantes :

$$\begin{cases} X_{k|k-1} = AX_{k-1|k-1} + w_{k-1} \\ P_{k|k-1} = AP_{k-1|k-1}A^T + Q_{k-1} \end{cases} \quad (2.2)$$

où  $X_{k|k-1}$  correspond à l'estimée a priori du vecteur d'état à l'instant  $k$ , calculée à l'instant  $k-1$  et  $P_{k|k-1}$  correspond à la matrice de covariance de l'erreur d'estimation.

b) une étape de correction qui améliore l'estimée a priori en tenant compte de la nouvelle mesure selon les équation suivantes :

$$\begin{cases} X_{k|k} = X_{k|k-1} + K_k(\Omega_k - h(X_{k|k-1})) \\ P_{k|k} = (I - K_k H_k)P_{k|k-1} \\ K_k = P_{k|k-1}H_k^T(R_k + H_k P_{k|k-1}H_k^T)^{-1} \end{cases} \quad (2.3)$$

avec  $K_k$  la matrice de gain de Kalman et  $H_k$  la matrice jacobienne de la fonction  $h$

$$H_k = \left. \frac{\partial h(X)}{\partial X} \right|_{X=X_{k|k-1}} \quad (2.4)$$

L'équation (2.4) correspond à la linéarisation de la fonction d'observation  $h$  (modèle de la caméra non linéaire) autour de la dernière estimée du vecteur d'état  $X_{k|k-1}$ .

***Remarque 2.7*** *Le filtre de Kalman est en fait la solution explicite de la minimisation d'un critère fonction de l'erreur entre les mesures et les estimées.*

Si une caméra de haute qualité est utilisée, le système est moins vulnérable aux perturbations comme les changements de luminosité. Il est alors possible de considérer que les paramètres représentant les bruits  $R_k$  et  $Q_k$  sont constants et connus au cours du temps. Cependant, pour des questions de coût, des caméras moins performantes sont généralement utilisées. Une adaptation des matrices de covariance d'erreur de processus  $Q_k$  et de mesure  $R_k$  en temps réel a été proposée [Bai98][Wir00][Lip04]. Dans [Lip04], une version adaptative de l'EKF, notée AEKF, est utilisée pour l'estimation d'objet 3D en mouvement. Les matrices  $Q_k$  et  $R_k$  sont mises à jour en tenant compte des  $N_q$  derniers échantillons et des  $N_r$  derniers échantillons ( $N_q$  pouvant être différent de  $N_r$ ).

### Remarque 2.8

- *Que ce soit pour l'EKF ou l'AEKF, les performances en terme d'estimation dépendent de la qualité du modèle de déplacement et du modèle de la caméra.*
- *Il est également important de noter qu'une estimée initiale du vecteur d'état  $X_0$  et de la matrice de covariance  $P_0$  sont requises pour pouvoir démarrer l'algorithme d'estimation. Une mauvaise initialisation de ces dernières peut engendrer des problèmes de convergence [Has05].*
- *Les bruits sont supposés blancs gaussiens non corrélés ce qui est rarement le cas en pratique.*
- *Le choix des matrices  $Q$  et  $R$  peut s'avérer difficile si l'on dispose de peu d'information.*
- *Des contraintes telles que la vitesse de déplacement de la personne ou des occultations dans l'image ne peuvent pas être prises en considération.*

Pour pallier ces limitations, nous proposons une solution alternative au filtre de Kalman étendu basée sur le principe de l'estimation sur horizon glissant. Sa formulation en terme de problème d'optimisation non linéaire est bien adaptée à la prise en compte de contraintes. Aucune hypothèse n'est formulée sur les bruits de mesures et de processus et aucune connaissance initiale n'est requise.

## 2.2 L'estimation de position sur horizon glissant

Nous allons présenter le principe de l'estimation sur horizon glissant, appelée en anglais Receding ou Moving Horizon Estimation (RHE ou MHE). Ensuite, nous étendrons son principe à l'estimation visuelle de position 3D. L'approche proposée est appelée Visual Receding Horizon Estimation (VRHE) puisque la minimisation se fait dans l'image.

### 2.2.1 Le principe de l'estimation d'état sur horizon glissant

L'estimateur sur horizon glissant a bénéficié des grandes avancées théoriques de la commande prédictive. Son principe est en effet le problème dual de la commande prédictive, aussi appelée commande optimale sur horizon glissant. On retrouve d'ailleurs les quatre points, communs à toute approche de commande prédictive.

Nous établissons, pour chacun de ces points, les correspondances (ou dualités) entre commande et estimation.

**La référence** (« ce que l'on veut »)

En commande, elle traduit les objectifs désirés en boucle fermée. C'est la trajectoire (fixe ou variable) à poursuivre.

En estimation, elle représente les mesures successives de la sortie du procédé.

**Le modèle** (« ce que l'on a »)

En commande, il sert à prédire, à partir d'une condition initiale à l'instant courant, le comportement futur du procédé sur un horizon de prédiction.

En estimation, il sert à estimer, à partir d'une condition initiale passée, le comportement passé du procédé sur un horizon d'estimation.

**Le critère** (aussi appelé fonction coût ou fonction objectif)

Il représente la minimisation de l'erreur entre « ce que l'on veut » et « ce que l'on a » sur un horizon fini glissant.

A chaque nouvelle mesure, ou mise à jour de l'information, l'horizon glisse d'une période d'échantillonnage.

En commande, l'horizon est un horizon de prédiction et la somme des erreurs est à minimiser par rapport à un vecteur de commandes.

En estimation, l'horizon est un horizon passé et la somme des erreurs est à minimiser par rapport au vecteur d'état en début d'horizon.

**La méthode de résolution**

Les procédés étant généralement non linéaires et soumis à des contraintes, le problème de minimisation est un problème d'optimisation non linéaire sous contraintes. Sa solution est itérative et obtenue grâce à un algorithme d'optimisation adéquat. Le temps de calcul est un facteur important pour une application temps réel.

Les différents blocs essentiels de l'estimation sur horizon glissant ayant été définis, nous nous intéressons maintenant à son principe illustré figures 2.1 et 2.2.

A l'instant courant  $k.T_e$  ( $T_e$  est la période d'échantillonnage), on effectue une mesure du procédé. Disposant des  $N_e$  mesures antérieures, on essaie de trouver la condition initiale « optimale »  $x[(k - N_e)T_e]$  qui minimise l'erreur entre les mesures du procédé et les sorties estimées du modèle sur l'horizon  $N_e$  (figure 2.1).

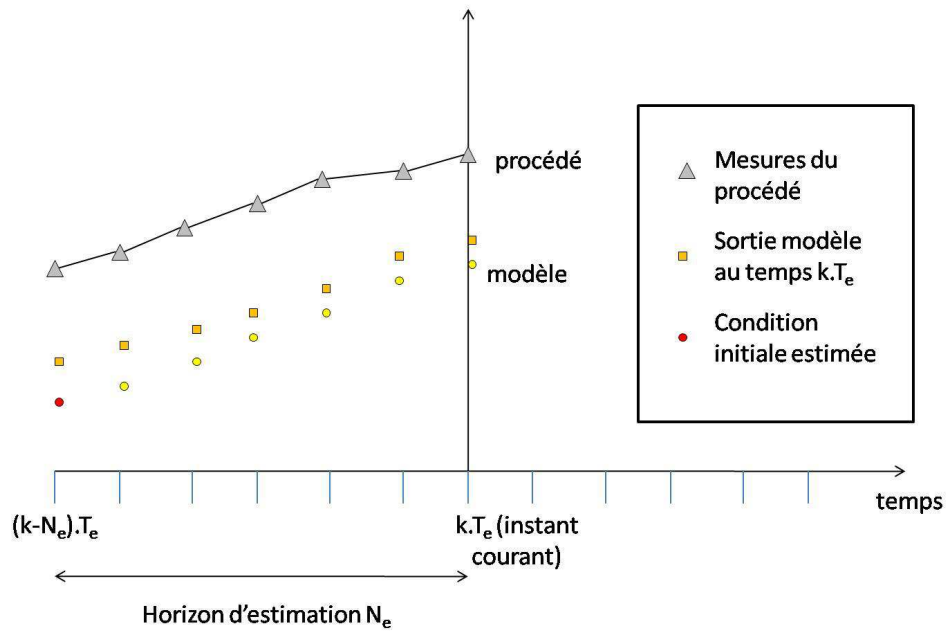


FIG. 2.1 – Principe de l'estimation d'état au temps  $k.T_e$ .

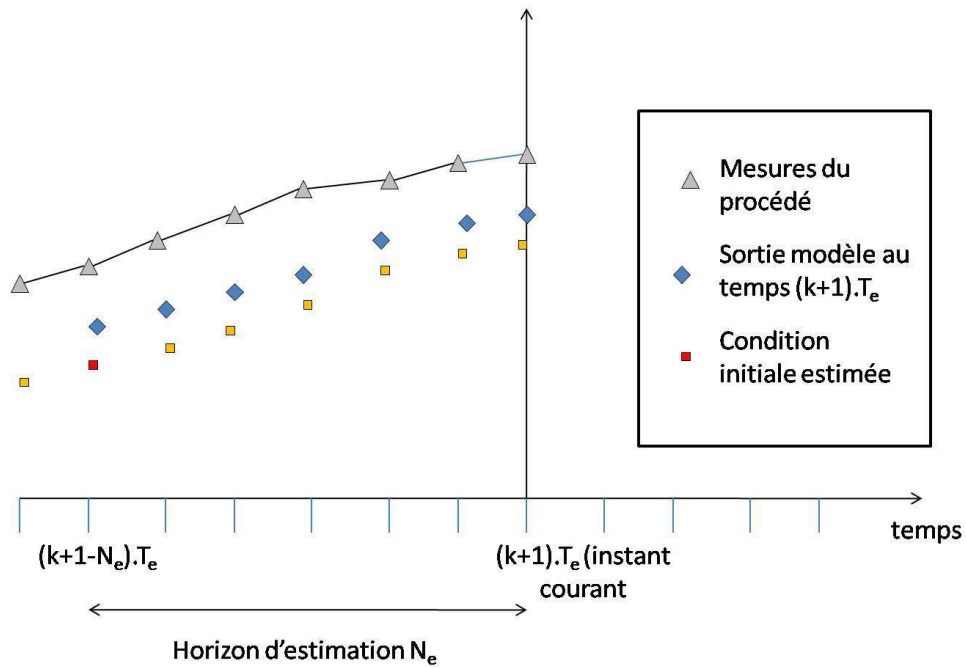


FIG. 2.2 – Principe de l'estimation d'état au temps  $(k+1).T_e$ .

Une fois le vecteur d'état  $x[(k-N_e)T_e]$  déterminé, on utilise le modèle en boucle ouverte pour estimer le vecteur d'état à l'instant  $k.T_e$ . Étant donné la présence inévitable de perturbations et d'incertitudes paramétriques sur le modèle, la procédure d'estimation est

recommencée à chaque nouvelle mesure. L’horizon d’estimation glisse alors d’une période d’échantillonnage (figure 2.2).

**Remarque 2.9** *L’horizon d’estimation  $N_e$  est de taille fixe et finie. Sa longueur est choisie en fonction des objectifs et des contraintes de l’application en gardant à l’esprit qu’une dimension importante entraîne un temps de calcul plus conséquent.*

Initialement, l’estimation sur horizon glissant a été appliquée dans le domaine du génie des procédés [Cou96][Tyl96][Hed07]. Les procédés ont des dynamiques relativement lentes et le temps de calcul n’est pas un facteur limitant. Ces différents exemples ont toutefois permis de montrer l’efficacité de l’approche dans la prise en compte explicite des contraintes. Dans [Has05], l’approche RHE est comparée à l’EKF (même modèle et mêmes mesures). La prise en compte de contraintes permet au RHE de converger tandis que l’EKF ne converge pas à coup sûr à cause d’une sensibilité vis à vis de l’initialisation de l’algorithme.

Actuellement, l’estimation sur horizon glissant est appliquée sur des systèmes rapides. Le temps de calcul a pu grandement être diminué grâce aux avancées récentes de la programmation non linéaire [Zav08].

Les raisons ci-dessus nous ont conforté dans l’idée d’utiliser l’estimation sur horizon glissant pour des tâches de localisation 3D à partir de l’image.

## 2.2.2 L’estimation visuelle sur horizon glissant

L’estimation visuelle sur horizon glissant est une extension de l’estimation sur horizon glissant à partir de données visuelles (informations 2D issues d’une image). Nous l’avons nommée en anglais Visual Receding Horizon Estimation (VRHE).

La problématique revient à déterminer la position 3D de la personne qui minimise, sur un horizon, l’écart entre les mesures et les estimées dans le plan image. Les estimées sont obtenues grâce à un modèle global décrit ci-dessous.

### 2.2.2.1 Le modèle global

Le modèle global, pièce maîtresse de l’approche, a un rôle essentiel dans le sens où il permet d’estimer l’évolution du système dans l’image sur l’horizon d’estimation. Ce modèle doit traduire le plus fidèlement possible le comportement du procédé observé. Le caractère global du modèle vient du fait qu’il combine à la fois le modèle de déplacement de la personne et le modèle de la caméra. En effet, les entrées du modèle global

correspondent aux déplacements de la personne tandis que les sorties sont le ou les points d'intérêt dans l'image.

### (a) Le modèle de la caméra

Afin d'exploiter au mieux les données délivrées par la caméra, il est nécessaire de connaître les changements de repère entre un repère lié à l'environnement  $R_e$  et le repère image  $R_{im}$  (figure 2.3), ainsi que les caractéristiques de la caméra. Une calibration de la caméra permet de déterminer ces paramètres.

**Remarque 2.10** *La calibration consiste à identifier les paramètres intrinsèques et extrinsèques de la caméra, à savoir sa distance focale, la dimension d'un pixel, les coordonnées du centre optique dans l'image ainsi que les paramètres de rotation et de translation pour passer du repère environnement au repère caméra.*

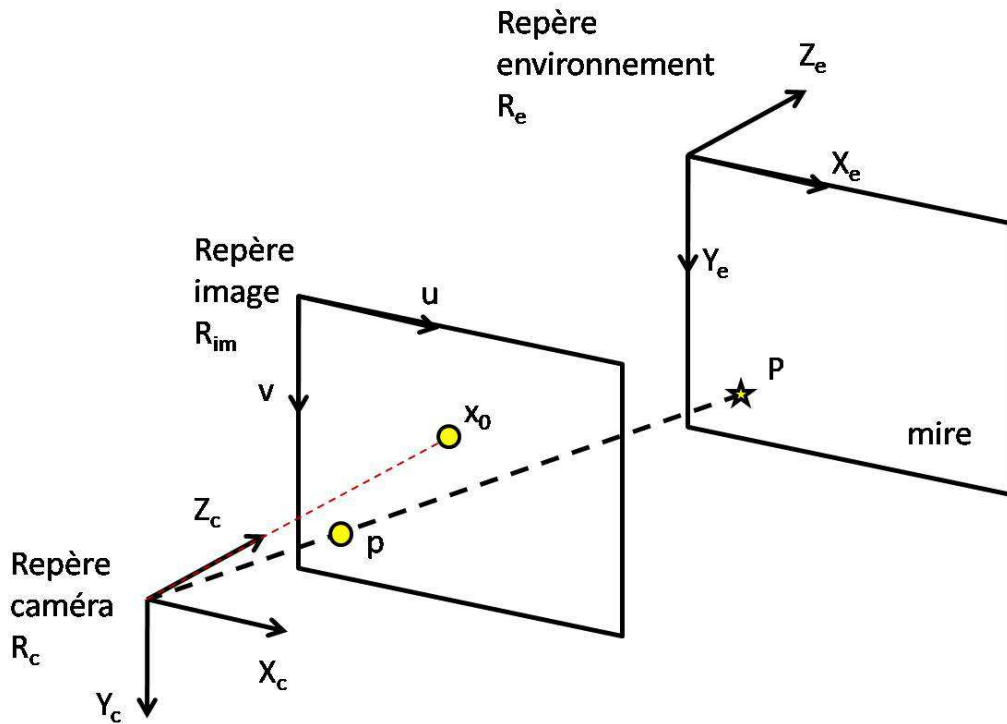


FIG. 2.3 – Représentation des repères définis lors de la calibration de la caméra.



La phase de projection dans l'image d'un point de l'espace 3D peut être décomposée en trois transformations successives :

1. La première consiste à exprimer les coordonnées d'un point 3D du repère environnement  $R_e$  dans le repère de la caméra  $R_c$  ;
2. La deuxième étape revient à utiliser le modèle de projection perspective afin de projeter le point 3D sur le plan image ;
3. La dernière étape détermine les coordonnées pixelliques du point dans l'image à l'aide des paramètres intrinsèques de la caméra.

### 1. Changement de repère

L'objectif est d'exprimer les coordonnées d'un point  $P$  de l'environnement observé, exprimées dans le repère  $R_e$  lié à une mire, dans le repère de la caméra  $R_c$ . Cela nécessite de connaître la position et l'orientation de la caméra par rapport à  $R_e$  afin de déterminer la matrice homogène  $T_{extrin}$  correspondant au changement de repère entre  $R_e$  et  $R_c$ . Si  $P$  a pour coordonnées homogènes  $P_e = (X_e, Y_e, Z_e, 1)^T$  dans le repère  $R_e$ , on obtient alors :

$$P_c = T_{extrin} P_e = \begin{pmatrix} R_{11} & R_{12} & R_{13} & T_x \\ R_{21} & R_{22} & R_{23} & T_y \\ R_{31} & R_{32} & R_{33} & T_z \\ 0 & 0 & 0 & 1 \end{pmatrix} P_e = \begin{pmatrix} R_{3 \times 3} & T_{3 \times 1} \\ 0_{1 \times 3} & 1 \end{pmatrix} P_e \quad (2.5)$$

avec  $P_c = (X_c, Y_c, Z_c, 1)^T$  le vecteur des coordonnées homogènes du point  $P$  dans le repère caméra  $R_c$ . La matrice  $R$  et le vecteur  $T$  représentent respectivement la matrice de rotation et le vecteur de translation entre les deux repères  $R_e$  et  $R_c$ .

### 2. Projection dans le plan image

L'étape suivante consiste à projeter le point  $P$  dans le plan image  $R_{im}$ . Ce plan, perpendiculaire au plan formé par  $X_c$  et  $Y_c$  du repère  $R_c$ , est situé à une distance  $f$  suivant  $Z_c$ ,  $f$  étant la distance focale de la caméra. Les coordonnées de  $P$  projeté sur le plan  $R_{im}$ , notées  $p = (x, y, z)^T$  sont obtenues via la relation :

$$p = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} f \frac{X_c}{Z_c} \\ f \frac{Y_c}{Z_c} \\ f \end{pmatrix} \quad (2.6)$$

Par commodité, puisque tous les points appartenant au plan image ont la même coordonnée suivant l'axe  $Z_c$ , le changement de variable suivant est généralement effectué :

$$\bar{p} = \frac{1}{f}p \quad (2.7)$$

Sous forme matricielle, on obtient ainsi :

$$Z_c \bar{p} = AP_c \quad (2.8)$$

avec

$$A = [I_3 \ 0_{3 \times 1}] \quad (2.9)$$

### 3. Passage aux coordonnées pixelliques

Enfin, on réalise le changement de coordonnées afin de passer du système métrique au système pixellique. Plusieurs paramètres obtenus lors de la phase de calibrage sont alors nécessaires :

- les dimensions d'un pixel suivant les axes  $X_c$  et  $Y_c$ , notées respectivement  $d_x$  et  $d_y$  ;
- les coordonnées  $(u_0, v_0)$  de la position du point  $x_0$  dans l'image.

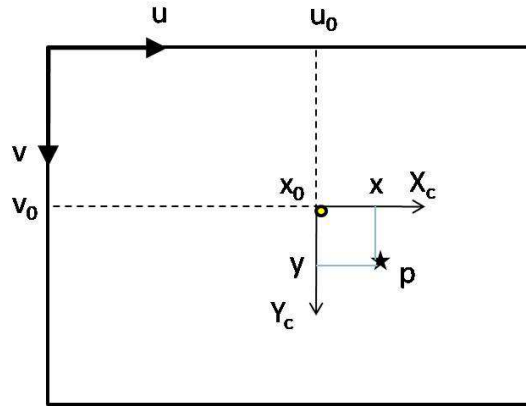


FIG. 2.4 – Plan image.

Les coordonnées pixelliques de  $p$ , notées  $u_p$  et  $v_p$ , s'expriment de la façon suivante (figure 2.4) :

$$\begin{pmatrix} u_p \\ v_p \end{pmatrix} = \begin{pmatrix} \frac{x}{d_x} + u_0 \\ \frac{y}{d_y} + v_0 \end{pmatrix} \quad (2.10)$$

Sous forme matricielle, on peut alors écrire :

$$\begin{pmatrix} u_p \\ v_p \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{d_x} & 0 & \frac{u_0}{f} \\ 0 & \frac{1}{d_y} & \frac{v_0}{f} \\ 0 & 0 & \frac{1}{f} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = Kp \quad (2.11)$$

En utilisant le changement de variable défini en (2.7), on obtient :

$$\begin{pmatrix} u_p \\ v_p \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{f}{d_x} & 0 & u_0 \\ 0 & \frac{f}{d_y} & v_0 \\ 0 & 0 & 1 \end{pmatrix} \bar{p} = T_{int} \bar{p} \quad (2.12)$$

La matrice  $T_{int}$  rassemble tous les paramètres intrinsèques de la caméra et peut se noter comme suit :

$$T_{int} = \begin{pmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \quad (2.13)$$

#### 4. Expression générale du modèle

La forme générale du modèle  $M_{cam}$  de la caméra s'obtient donc en regroupant les équations (2.5), (2.9) et (2.13) et se présente de la manière suivante :

$$\begin{pmatrix} u_p \\ v_p \\ 1 \end{pmatrix} = T_{int} A T_{extrin} \begin{pmatrix} X_e \\ Y_e \\ Z_e \\ 1 \end{pmatrix} = M_{cam} P_e \quad (2.14)$$

avec  $T_{int}$  la matrice des paramètres intrinsèques,  $A$  la matrice de projection normalisée et  $T_{extrin}$  la matrice de transformation homogène entre les repères  $R_e$  et  $R_c$ .

**Remarque 2.11** *Le modèle défini en (2.14) est un modèle simplifié amplement suffisant pour la majorité des applications en vision. Il peut être nécessaire, dans certains cas, d'introduire des termes correctifs afin de compenser les éventuelles distorsions dans l'image.*

**Remarque 2.12** *Nous utilisons, dans les programmes, la matrice  $M_{cam}$  obtenue directement lors de la calibration. Pour chaque campagne de tests, une détermination du modèle de la caméra a été réalisée, étant donné que l'environnement et la position de la caméra pouvaient changer à chaque campagne.*

## (b) Le modèle de déplacement

La locomotion humaine ne suit pas des règles précises et possède même un caractère imprévisible, complexifiant davantage sa modélisation. Une première approche [Bru09b], très simpliste, a consisté à modéliser le déplacement humain à l'aide d'un intégrateur pur, qui après une discrétisation du premier ordre, a conduit à un modèle linéaire. Puis, nous avons recherché un modèle plus représentatif de la mécanique de marche d'un humain [Bru09a]. Développé pour la robotique humanoïde, un modèle non holonome de locomotion humaine a retenu notre attention [Are06a][Are06b].

Si l'on s'intéresse à la manière dont se déplace l'homme, il apparaît effectivement qu'il se déplace rarement latéralement, mouvement qui est également impossible pour un robot mobile non holonome comme une voiture. Des contraintes de natures différentes restreignent les possibilités de trajectoires du déplacement humain, comme des contraintes liées à l'anatomie (bassin) et à la mécanique du corps humain. Dans [Are06b], les auteurs montrent que le corps humain est comparable à une voiture muni d'un volant. Ainsi, sous l'hypothèse d'un déplacement sans glissement, le modèle cinématique de la locomotion humaine peut s'écrire de la manière suivante :

$$\begin{pmatrix} \dot{x}_h(t) \\ \dot{y}_h(t) \\ \dot{\theta}_h(t) \end{pmatrix} = \begin{pmatrix} \cos \theta_h(t) \\ \sin \theta_h(t) \\ 0 \end{pmatrix} v(t) + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \omega(t) = f(p_h(t), u(t)) \quad (2.15)$$

où  $p_h(t) = (x_h(t), y_h(t), \theta_h(t))^T$  correspond à la posture, position et orientation, de la personne dans le repère  $R_h$  (figure 2.5) et  $u(t) = (v(t), \omega(t))$  est le vecteur d'entrées composé respectivement des vitesses linéaire et angulaire.

**Remarque 2.13** *Le vecteur  $P_h = (x_h, y_h, z_h, 1)^T$ , vecteur des coordonnées homogènes de la personne dans le repère  $R_h$ , est obtenu à partir des deux premières composantes du vecteur de posture  $p_h$  (2.15). La troisième coordonnée, i.e.  $z_h$ , est comme on l'a précisé précédemment, supposée constante (hypothèse du déplacement sur un plan). On peut donc écrire la relation suivante :*

$$P_h(t) = g(p_h(t)) \quad (2.16)$$

## (c) Transformation homographique

Le modèle de déplacement est défini dans le repère 3D  $R_h$  lié à la surface de la pièce. Une transformation homographie planaire est nécessaire pour passer du repère image  $R_{im}$

vers le repère  $R_h$ . Cette transformation permet en effet, sous l'hypothèse d'un sol plat, de calculer les coordonnées d'un point dans un repère lié à l'environnement à partir de ses coordonnées dans l'image.

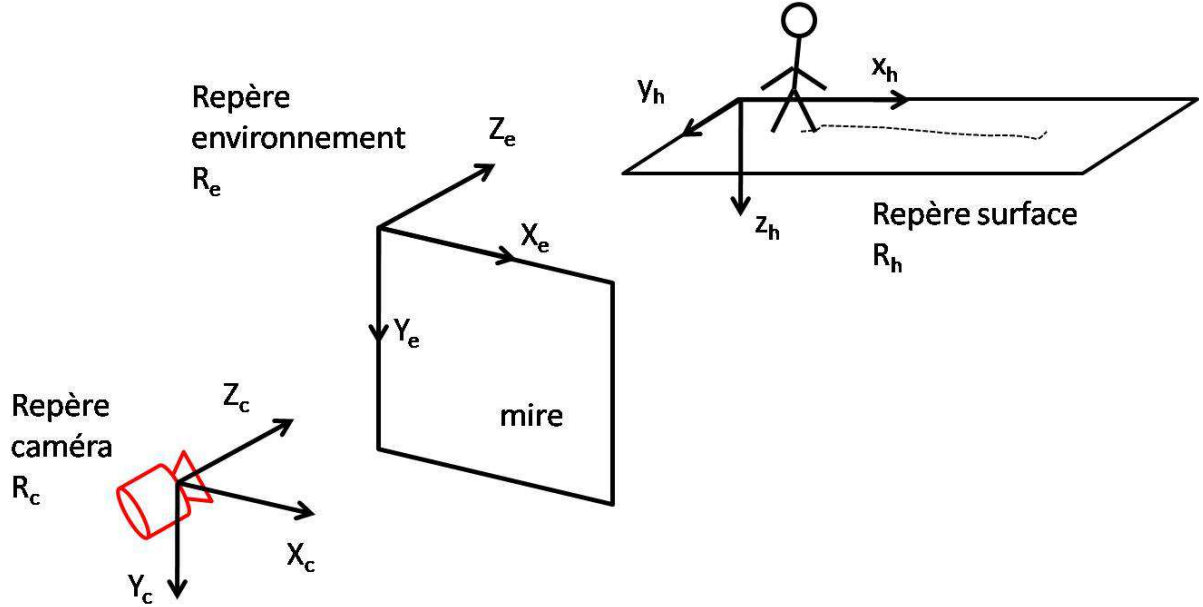


FIG. 2.5 – Estimation de position à l'aide d'une caméra perspective.

Le repère surface  $R_h$  utilisé lors de cette homographie planaire est différent du repère environnement  $R_e$  défini lors de la phase de calibration de la caméra (figure 2.5). Une transformation entre ces deux repères est nécessaire afin de déterminer les positions dans l'image à partir des positions dans l'environnement, positions issues du modèle de déplacement. Cette transformation se présente sous la forme d'une matrice,  $M_{trans}$ , qui est une matrice de transformation classique, combinaison de rotations et de translation :

$$M_{trans} = \begin{pmatrix} 1 & 0 & 0 & X_r \\ 0 & 0 & 1 & Y_r \\ 0 & -1 & 0 & Z_r \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (2.17)$$

$$P_e(t) = \begin{pmatrix} X_e \\ Y_e \\ Z_e \\ 1 \end{pmatrix} = M_{trans} \begin{pmatrix} x_h \\ y_h \\ z_h \\ 1 \end{pmatrix} = tr(P_h(t)). \quad (2.18)$$

où :

- $P_h = (x_h, y_h, z_h, 1)^T$  sont les coordonnées homogènes de l'humain dans le repère  $R_h$ ,
- $P_e = (X_e, Y_e, Z_e, 1)^T$  sont les coordonnées homogènes de l'humain dans le repère environnement  $R_e$  utilisé lors de la calibration,
- $(X_r, Y_r, Z_r)$  sont les coordonnées de l'origine du repère  $R_h$  par rapport au repère  $R_e$ .

#### (d) Le modèle global

Il s'obtient en combinant le modèle de la caméra (2.14), le modèle de déplacement (2.15) (2.16) et la matrice de transformation  $M_{trans}$  (2.18) :

$$(M_g) \begin{cases} \dot{p}_h(t) = f(p_h(t), u(t)) \\ y_{mod}(t) = M_{cam} \cdot tr(g(p_h(t))) = h(p_h(t)) \end{cases} \quad (2.19)$$

où :

- $u(t) = (v(t), \omega(t))^T$  est le vecteur d'entrées,
- $p_h(t) = (x_h(t), y_h(t), \theta_h(t))^T$  est le vecteur de posture de la personne dans le repère surface  $R_h$ ,
- $y_{mod} = (u_{mod}, v_{mod}, 1)^T$  est le vecteur de sortie, c'est à dire la projection dans le plan image de la position des pieds.

Nous disposons alors d'un modèle global  $(M_g)$  (2.19), qui est un modèle non linéaire traduisant à la fois le déplacement de la personne et la modélisation du dispositif visuel ainsi que les différents changements de repères nécessaires. Un des avantages de ce modèle réside dans sa facilité d'adaptation si le dispositif visuel ou le modèle de déplacement est différent.

#### 2.2.2.2 La formulation mathématique du VRHE

Le critère d'optimisation est fonction des écarts entre les sorties mesurées et les sorties estimées par le modèle sur l'horizon d'estimation passé.

Cette fonction coût s'écrit en temps discret de la manière suivante :

$$\min_{p_h((k-N_e)T_e)} J(p_h) = \sum_{j=(k-N_e)T_e}^{kT_e} [y_{mes}(j) - y_{mod}(j)]^T Q [y_{mes}(j) - y_{mod}(j)], \quad (2.20)$$

avec :

$$\begin{cases} \dot{p}_h(t) = f(p_h(t), u(t)) \\ y_{mod}(j.T_e) = h(p_h(j.T_e)) \end{cases}$$

où :

- $p_h(t) = (x_h(t), y_h(t), \theta_h(t))^T$  et  $y_{mod} = (u_{mod}, v_{mod}, 1)^T$  sont respectivement l'état et la sortie du modèle global ;
- $u(t) = (v(t), \omega(t))^T$  est le vecteur d'entrée du modèle global ;
- $y_{mes} = (u_p, v_p, 1)^T$  correspond au vecteur des mesures dans le repère image ;
- $N_e$  est l'horizon d'estimation ;
- $Q$  est une matrice de pondération symétrique définie positive.

### Les contraintes

Un des atouts de l'estimation visuelle sur horizon glissant est la possibilité de prendre en compte explicitement des contraintes.

Dans le cas de la localisation 3D d'une personne, on distingue deux types de contraintes :

- Les contraintes portant sur le système observé. Pour une application de localisation humaine, le déplacement comporte des limitations inhérentes à la mécanique du corps humain en terme de vitesse et d'accélération ;
- Les contraintes dans l'image. Les estimations sont bornées par le cadre de l'image et la présence d'obstacles constitue une contrainte en terme de visibilité.

Les contraintes portant sur le système sont des contraintes sur les entrées. Elles peuvent s'exprimer par des limitations en amplitude.

Par exemple :

$$v \leq v_{max}$$

où  $v_{max}$  représente le seuil maximal de déplacement admissible.

Les contraintes sur l'image sont très importantes. L'intérêt pour une application visuelle est également de s'assurer que les estimées restent dans le champ de vision de la caméra. Pour des images de taille  $M \times N$  pixels, cette contrainte peut se traduire de la manière suivante :

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} \leq y_{mod} \leq \begin{pmatrix} M \\ N \end{pmatrix} \quad (2.21)$$

Dans le cas d'occultation dans l'image, on peut définir cette limitation comme une contrainte de visibilité sous forme d'inégalités.

$$\text{Si } u_{limb} < u_p < u_{limh} \text{ alors } v_p \leq v_{limh} \text{ ou } v_p \geq v_{limb}$$

Ces contraintes s'ajoutent facilement au problème d'optimisation (2.20) qui devient un problème d'optimisation non linéaire sous contraintes.

Pour la résolution de ce problème, il existe de nombreux algorithmes d'optimisation. Leur choix dépend du temps de calcul disponible entre deux instants d'échantillonnage (le temps imparti au traitement d'images ayant été enlevé).

Cette approche, inspirée de l'estimation sur horizon glissant transpose la problématique de localisation 3D d'une personne en un problème d'optimisation non linéaire dans le plan image avec la possibilité de prendre explicitement en compte des contraintes. C'est principalement en cela qu'elle se différencie des approches classiquement utilisées comme le filtre de Kalman étendu qui n'offre pas cette possibilité. De plus, elle utilise pour l'étape de prédiction un modèle global non linéaire (pas de linéarisation autour d'un point), offrant ainsi une meilleure représentation des phénomènes observés. Ce modèle global, véritable clé de voûte pour l'estimation de l'évolution du système dans l'image, combine à la fois un modèle de déplacement de la personne et un modèle du dispositif visuel. Il convient de préciser pour conclure que l'approche n'impose aucun système spécifique de vision ou de déplacement, apportant ainsi flexibilité et adaptabilité à l'approche développée.

## 2.3 Application du VRHE à la localisation 3D d'une personne

Après avoir décrit les conditions expérimentales, des résultats expérimentaux permettent d'illustrer la faisabilité de la méthode et les possibilités de l'approche en présence d'occultations dans l'image.

### 2.3.1 Les conditions expérimentales

- On suppose que la personne se déplace sur un sol plat.

**Remarque 2.14** *Cette hypothèse peu restrictive se justifie par le fait qu'en maintien à domicile, les changements de niveaux sont généralement évités afin de ne pas gêner le déplacement d'une personne âgée et/ou handicapée et de ne pas accentuer également les risques de chutes.*

- Les séquences d'enregistrement réalisées traduisent des situations de la vie courante comme des déplacements dans une pièce meublée. Deux catégories de séquences ont été



constituées. Dans la première catégorie, il n’y a aucun obstacle susceptible de gêner la visibilité du dispositif de vision dans l’environnement surveillé. Notre but est d’évaluer notre approche sur des scénarios simples et de vérifier la pertinence de nos choix notamment en terme de modélisation du déplacement humain. L’approche développée est comparée avec le filtre de Kalman étendu. Dans un second temps, des scénarios ont été réalisés en présence d’obstacles représentant, par exemple, une table au milieu d’une pièce. L’approche développée est alors testée en présence de contraintes afin d’illustrer les potentialités de l’estimation sur horizon glissant.

- L’ensemble des séquences a été enregistré à l’aide d’une seule caméra perspective, ayant une résolution de  $320 \times 240$  pixels, positionnée en hauteur (environ 2 mètres) dans le coin d’une pièce. Cela permet d’avoir une couverture quasi totale de l’environnement surveillé et de se rapprocher au mieux des futures conditions d’installation du dispositif CAPTHOM (figure 2.6).



FIG. 2.6 – Environnement vu par la caméra perspective.

***Remarque 2.15*** Lors de nos expérimentations, les images ont été conservées afin d’illustrer les résultats sur le VRHE mais nous tenons à préciser que seules certaines données issues du traitement d’images sont nécessaires au fonctionnement du VRHE et qu’en pratique, aucune image ne sera conservée ou transmise par le dispositif CAPTHOM. Ceci bien entendu pour le respect de l’intimité des personnes. La plateforme du GIS MADONAH n’étant pas équipée de caméra, les expérimentations ont eu lieu en laboratoire.

- D'un point de vue logiciel, l'algorithme de traitement d'images a été développé en C++ sous Visual Studio avec la bibliothèque OpenCV. L'approche VRHE a été implémentée sous Matlab (fonction `fmincon` pour l'optimisation non linéaire sous contraintes). L'ensemble des calculs ont été réalisés hors ligne à partir des données issues de l'algorithme de traitement d'images.

- L'algorithme de traitement d'images a été développé par [Ben09b]. Cet algorithme se classe dans la catégorie des méthodes basées sur l'utilisation d'un apprentissage supervisé et d'un balayage de l'image. La constitution d'une base d'images est un pré-requis de ces méthodes, base d'images rassemblant à la fois des représentations de personnes (le plus souvent debout) dans des environnements variés et des images ne comportant pas d'humains. L'objectif est, à travers un algorithme d'apprentissage, de déterminer des classificateurs qui permettent de caractériser l'entité détectée dans chaque fenêtre comme étant humaine ou non. La classification de la cible en tant qu'humain constitue la dernière étape de l'algorithme de traitement d'images.



(a)



(b)



(c)



(d)

FIG. 2.7 – Les différentes phases de l'algorithme de traitement d'images.

La figure 2.7 illustre les principales phases successives du traitement d'image.

- La première phase consiste à capturer quelques images de la pièce, vide de toute présence, afin de disposer d'un modèle de l'arrière plan auquel toutes les futures captures seront comparées (figure 2.7 (a)).

- Dans un second temps, chaque nouvelle image provenant de la caméra perspective est comparée avec le modèle d'arrière plan afin de détecter les changements potentiels. Basée sur l'utilisation de la distance de Mahalanobis entre les pixels des deux images et d'un seuil, cette comparaison permet de réduire l'espace de recherche sur lequel est appliqué le classifieur, évitant ainsi de devoir balayer l'ensemble de l'image et optimisant le temps de calcul de l'algorithme (figure 2.7 (b)). A chaque itération, le modèle de l'arrière plan est mis à jour : mise à jour de chaque pixel par un filtre moyennneur temporel, si un objet statique est caractérisé comme étant non humain tous les pixels de l'avant plan correspondant sont inclus dans le modèle de l'arrière plan et en cas de changement brusque de la luminosité, une réinitialisation du modèle est effectuée.

**Remarque 2.16** *La distance de Mahalanobis est une mesure de distance qui permet de déterminer la similarité entre deux séries de données [Mah36]. En traitement d'images, le calcul de cette distance pour chaque pixel prend en compte la distribution de couleur du pixel, la moyenne et la matrice de covariance.*

- Un suivi de points d'intérêt est ensuite opéré (figure 2.7 (c)) afin d'être capable de gérer la présence de plusieurs personnes dans la pièce et de connecter chaque changement de pixel avec le bon objet. Cette étape permet également de prévoir la position future de la personne dans l'image et de gérer au mieux des situations de croisement de personnes ou d'apparition d'une nouvelle entité dans la pièce. Chaque cible doit garder en permanence le même identifiant.

- Ce n'est qu'après ces étapes que la discrimination de la cible entre une entité humaine ou non humaine est effectuée (figure 2.7 (d)).

Afin d'illustrer le résultat de la détection, la personne est encadrée par un rectangle dont nous disposons des coordonnées. La position des pieds de la personne dans l'image peut être assimilée au milieu du côté inférieur de la boîte englobante. On obtient alors une localisation approximative, dans le plan image, de la personne détectée :

$$y_{mes} = (u_{mes}, v_{mes}).$$

- Après calibration, on obtient le modèle de la caméra  $M_{cam}$  (2.14) :

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} 329.7599 & 35.8226 & 206.2159 & 158.7368 \\ -16.2052 & 367.7602 & 5.1678 & 81.8475 \\ -0.1535 & 0.1721 & 0.993 & 0.92 \end{pmatrix} \begin{pmatrix} X_e \\ Y_e \\ Z_e \\ 1 \end{pmatrix} \quad (2.22)$$

• Transformation homographique. La matrice d'homographie  $M_{hom}$  relie les pixels de l'image à leurs positions sur la surface réelle. En supposant un déplacement sur un sol plat, la détermination de cette matrice nécessite l'utilisation de quatre points, matérialisés par un repère rectangulaire ( $0.92 \times 0.59$ m) dont on connaît les coordonnées réelles et pixelliques (figure 2.8).

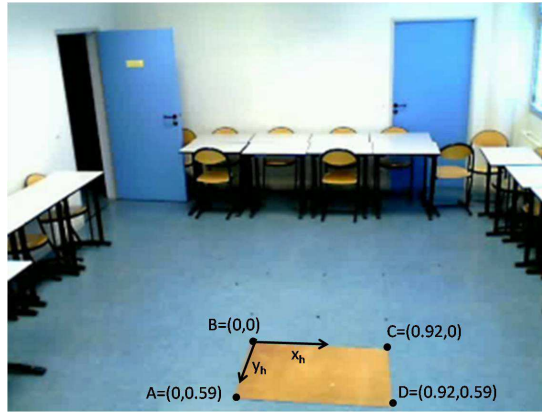


FIG. 2.8 – Représentation du repère  $R_h$  utilisé pour l'homographie.

**Remarque 2.17** La méthode de détermination de la matrice  $M_{hom}$  est détaillée dans l'annexe A.

### 2.3.2 Expérience sans contrainte

Les premières simulations présentées ci-dessous ont pour objectif de montrer la faisabilité de la méthode sur des exemples de scénarios simples où aucun obstacle n'est présent dans le champ de vision de la caméra.

#### Conditions de simulation

- Les trajectoires réelles dans le repère  $R_h$  ont été relevées afin de les comparer aux estimations obtenues, mais elles ne sont pas utilisées dans les calculs ;
- La période d'échantillonnage  $T_e$  est égale à 70ms ;
- La matrice  $Q$  définie dans l'équation (2.20) est égale à la matrice identité ;

- Le problème d'optimisation non linéaire est résolu sous Matlab par la fonction *fminunc*.

### 2.3.2.1 Comparaison entre le modèle déplacement linéaire et le modèle non holonome

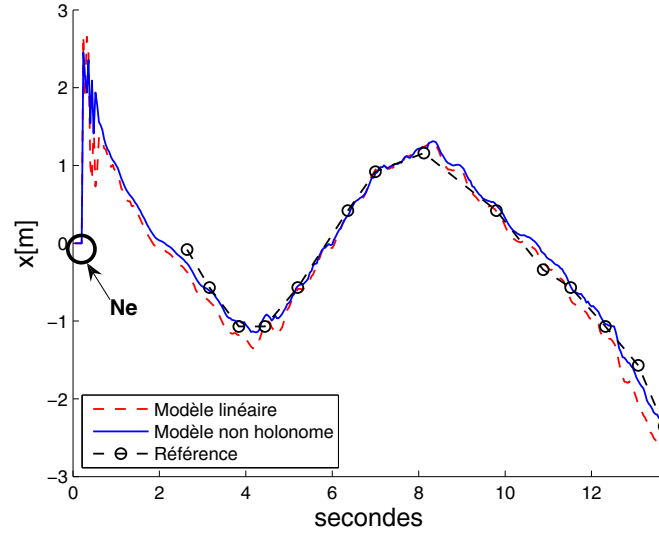


FIG. 2.9 – Comparaison entre les modèles linéaire et non holonome suivant l'axe  $x$  dans le repère  $R_h$ .

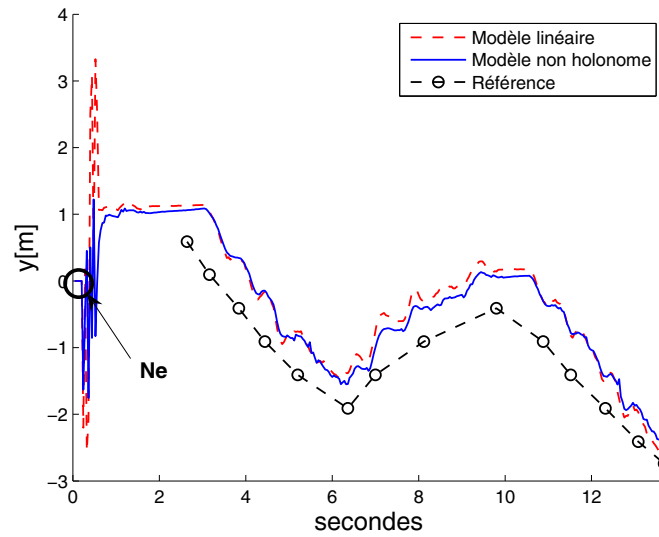


FIG. 2.10 – Comparaison entre les modèles linéaire et non holonome suivant l'axe  $y$  dans le repère  $R_h$ .

**Remarque 2.18** *L'estimation ne commence que lorsque l'algorithme dispose de  $N_e$  mesures.*

Comme nous l'avons précisé lors de la description du modèle de déplacement, deux modèles de locomotion ont été étudiés. Les figures 2.9 et 2.10 montrent que les deux modélisations offrent de bonnes performances en terme de convergence. Les estimations obtenues à l'aide du modèle de locomotion non holonome sont légèrement meilleures que celle obtenues avec le modèle linéaire, notamment suivant l'axe  $y_h$ . L'ensemble des simulations présentées par la suite utilise le modèle de locomotion non holonome.

### 2.3.2.2 Influence de l'horizon d'estimation $N_e$

L'objectif est d'étudier l'influence de l'horizon d'estimation sur la qualité de l'estimation et sur le temps de calcul. Pour cela, différentes trajectoires de déplacement ont été testées. L'horizon d'estimation  $N_e$  est égal successivement à 5, 10 et 15. La posture initiale de la personne est donnée par  $p_h = (0.08, 0.59, 0)^T$  pour le trajet N°1 et par  $p_h = (-2.35, -2, 74, 0)^T$  pour le trajet N°2.

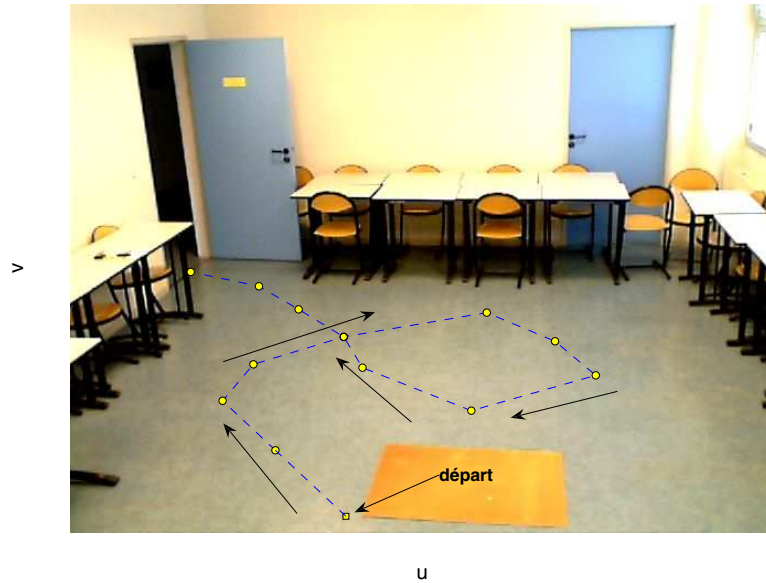


FIG. 2.11 – Représentation du trajet N°1.

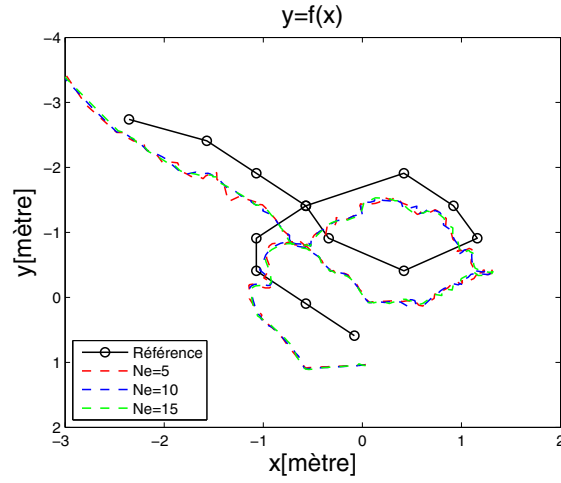
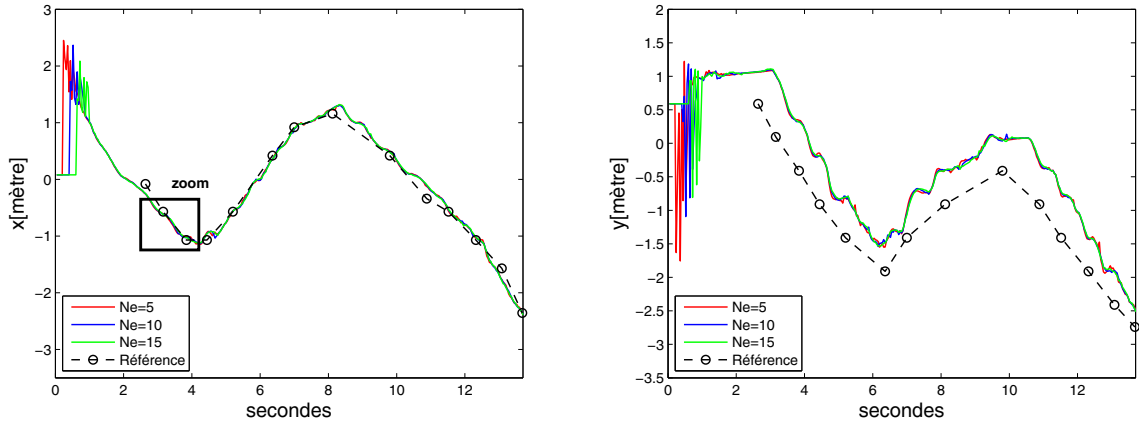
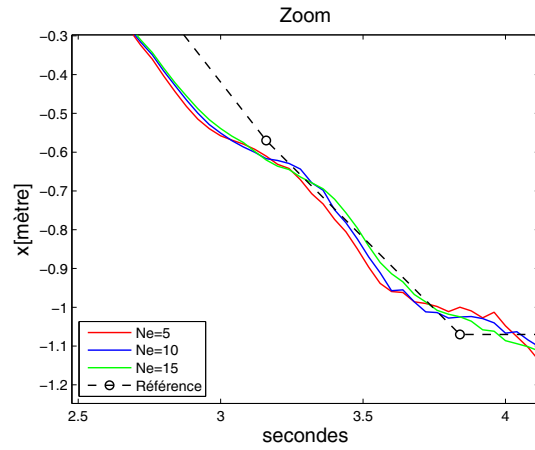


FIG. 2.12 – Estimation du trajet 1 pour  $N_e = 5, 10, 15$ .



(a)

(b)



(c)

FIG. 2.13 – Les positions en x et y dans  $R_h$  pour le trajet N°1 avec  $N_e = 5, 10, 15$ .

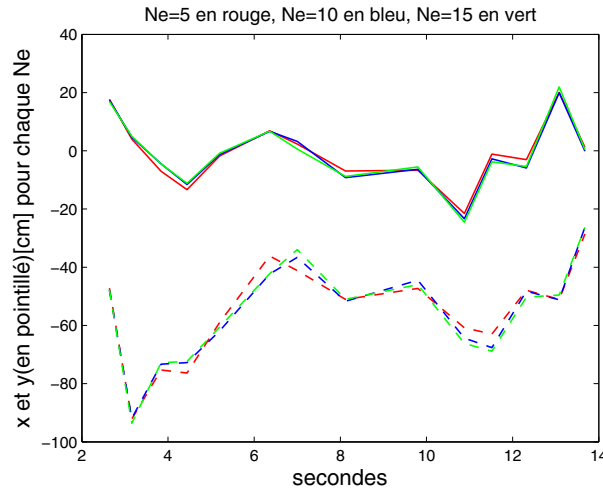


FIG. 2.14 – Erreur pour le trajet N°1 dans le repère  $R_h$  avec  $N_e = 5, 10, 15$ .

Les figures 2.12, 2.13, 2.14 montrent que pour le trajet N°1, quelle que soit la taille de l'horizon d'estimation  $N_e$ , les résultats sont presque identiques. La convergence est obtenue en  $x_h$  (figure 2.13 (c)) et on constate un léger décalage constant d'environ 50cm pour les estimations suivant  $y_h$  (figure 2.13 (b)).

**Remarque 2.19** Ce décalage peut s'expliquer par le reflet de la personne sur le sol qui est englobé dans le résultat de détection de l'algorithme de traitement d'images. Le bas de la boîte englobante se trouve alors à une position plus basse que les pieds dans l'image.

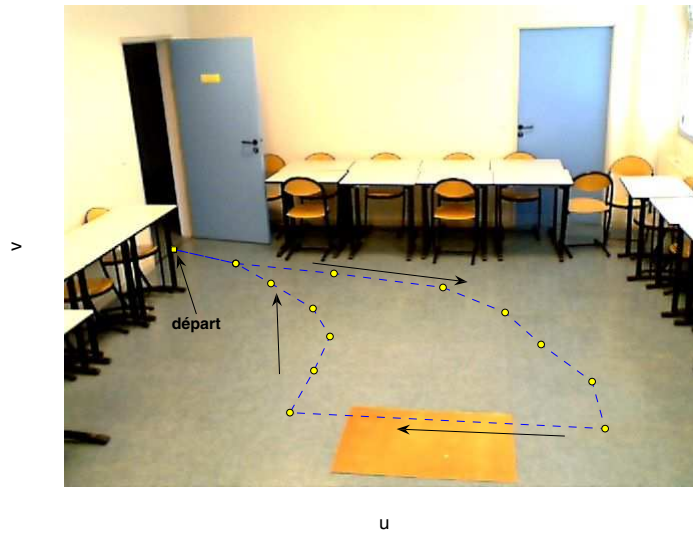


FIG. 2.15 – Représentation du trajet N°2.



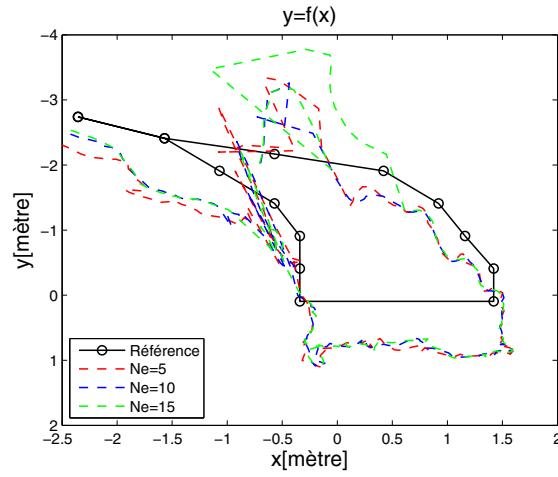
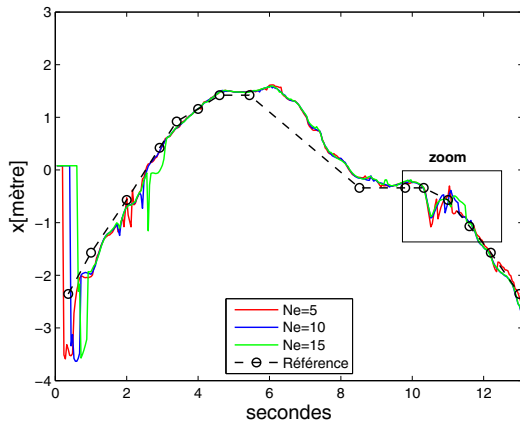
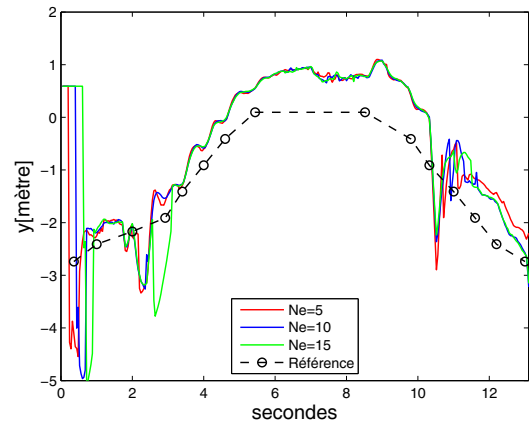


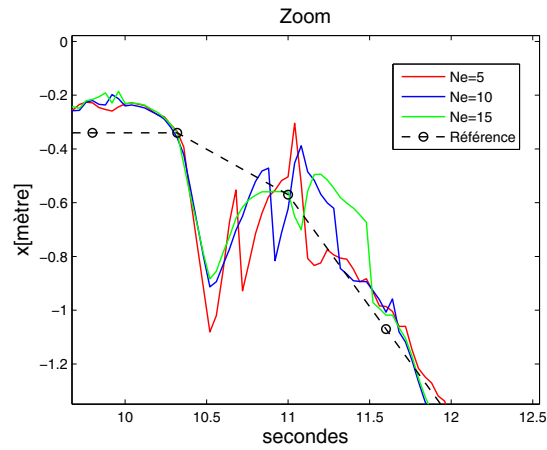
FIG. 2.16 – Estimation du trajet 2 pour  $N_e = 5, 10, 15$ .



(a)



(b)



(c)

FIG. 2.17 – Les positions en x et y dans  $R_h$  pour le trajet N°2 avec  $N_e = 5, 10, 15$ .

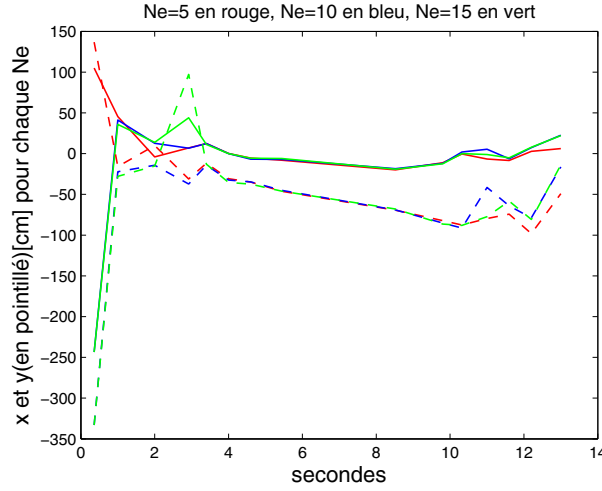


FIG. 2.18 – Erreur pour le trajet N°2 dans le repère  $R_h$  avec  $N_e = 5, 10, 15$ .

Pour le trajet N°2, les trajectoires estimées avec  $N_e = 5, 10, 15$  convergent vers la trajectoire réelle (figures 2.16, 2.17 et 2.18). Les fortes oscillations observées sur la figure 2.17 (b) s'expliquent par des brusques variations de la boîte englobante dues à une mauvaise soustraction de l'arrière plan pendant quelques pas d'échantillonnage. Les données utilisées sont brutes, sans filtrage préalable. Ces variations, selon l'axe  $y_h$ , entraînent de légères variations selon l'axe  $x_h$ . Lors de ces perturbations, un horizon d'estimation  $N_e = 5$  permet une convergence plus rapide vers la trajectoire réelle (figure 2.17 (c)).

**Remarque 2.20** *Le choix de  $N_e$  doit satisfaire un compromis entre qualité de l'estimation et temps de calcul. Un court horizon d'estimation est peu gourmand en temps de calcul, converge rapidement vers la trajectoire réelle et semble moins perturbé par des variations brusques. A l'inverse un horizon d'estimation grand va nécessiter un temps de convergence et un temps de calcul légèrement plus long. Les mêmes conclusions ont été obtenues pour des trajets différents.*

Afin de vérifier la faisabilité de la méthode pour une application temps réel, nous avons calculé, à l'aide de plusieurs simulations, la moyenne du temps de calcul nécessaire à la résolution du problème d'optimisation en fonction de la taille de l'horizon d'estimation (tableau 2.1).

L'horizon d'estimation  $N_e = 5$  semble être la meilleure solution, valeur que l'on utilise donc dans la suite de nos simulations.

$N_e$	Temps de calcul par itération (ms/iter)
$N_e = 5$	11.72 ms
$N_e = 10$	13.82 ms
$N_e = 15$	15.70 ms

TAB. 2.1 – Temps de calcul en fonction de l’horizon d’estimation.

### 2.3.2.3 Comparaison avec le filtre de Kalman étendu (EKF)

L’objectif de cette partie est de comparer l’approche VRHE avec le filtre de Kalman étendu, ce dernier étant l’estimateur le plus couramment utilisé en estimation visuelle de position. Les conditions expérimentales restent inchangées ( $N_e = 5$ ).

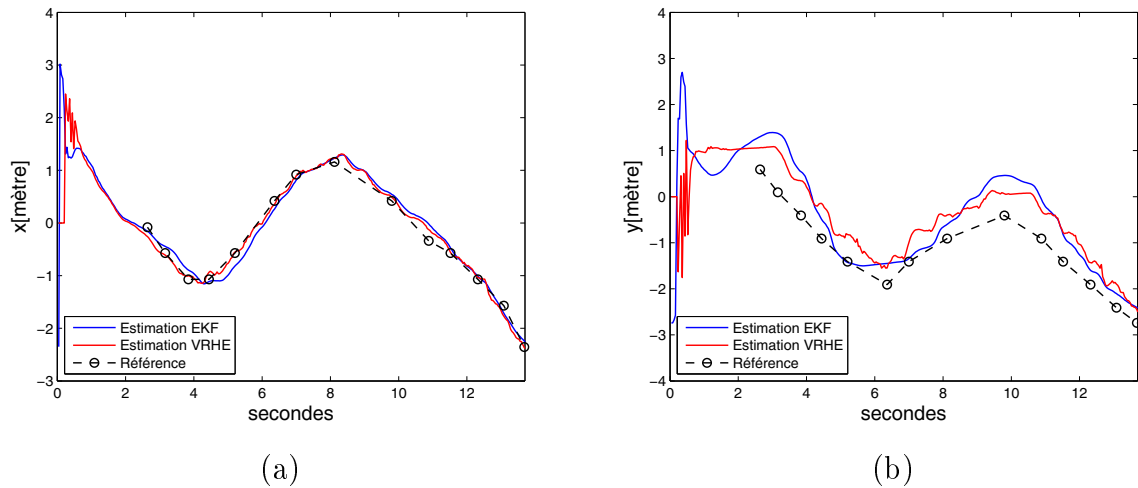


FIG. 2.19 – Comparaison des estimations de l’EKF et du VRHE suivant les axes  $x$  et  $y$  du repère  $R_h$ .

La figure 2.19 montre que le VRHE offre des performances équivalentes à l’EKF concernant l’estimation suivant l’axe des  $x$  du repère  $R_h$  et semble plus performant pour l’estimation de la position suivant l’axe des  $y$ . Au niveau temps de calcul, sur un même nombre d’échantillons, l’EKF est plus rapide que le VRHE, résultat logique dans la mesure où il ne travaille qu’avec la dernière mesure et non avec un horizon passé. Le VRHE gagne donc en performance ce qu’il perd en temps de calcul et se présente donc comme une bonne alternative à l’EKF.

### 2.3.3 Expérience avec contrainte

En présence d'un obstacle dans l'environnement surveillé, des problèmes d'occultations surviennent dès que la personne passe derrière l'obstacle. L'obstacle masque partiellement la personne. L'algorithme de traitement d'images détermine alors une boîte englobante plus petite. Étant donnée que la mesure des coordonnées de la personne dans le repère environnement est effectuée à partir des coordonnées du milieu du côté inférieur de la boîte englobante, cette variation importante des dimensions de la boîte conduit à des coordonnées en  $x$  et en  $y$  aberrantes par rapport aux précédentes estimations.



(a)



(b)



(c)



(d)

FIG. 2.20 – Influence d'un obstacle sur le résultat de détection de présence.

La figure 2.20 illustre le phénomène d’occultation et montre clairement les variations de taille de la boîte englobante. La vignette (a) illustre le champ de vision de la caméra ainsi que l’obstacle situé au milieu de la pièce. Dans la vignette (b), on constate que la détection de présence humaine est correctement réalisée. Dans la vignette (c), la boîte englobe uniquement la partie supérieure du corps. C’est un résultat prévisible dans la mesure où, lors de la phase de soustraction de l’arrière plan, aucun changement n’est détecté pour les pixels représentant l’obstacle. Sur la vignette (d), dès que la personne est de nouveau entièrement visible, la détection se fait de nouveau correctement.

**Remarque 2.21** *On constate sur la vignette (d) de la figure 2.20 que le reflet de la personne dans la vitre est aussi détecté et caractérisé comme étant une autre personne présente dans la pièce. Un détecteur IRP permet de lever cette ambiguïté sur le nombre de personnes.*

L’occultation a donc une influence importante sur l’estimation de position suivant  $y_h$  comme le montre la figure 2.21. Si l’on se fie aux mesures, la personne se serait déplacée d’environ 3m suivant l’axe  $y_h$  en quelques millisecondes. Pour pallier cette difficulté, nous allons considérer une contrainte de déplacement admissible sur les mouvements de la personne, en considérant que ses déplacements selon  $x_h$  et  $y_h$  sont limités en amplitude. L’écart en position suivant chaque axe ne peut pas dépasser une valeur seuil, fonction de la moyenne des précédents déplacements.

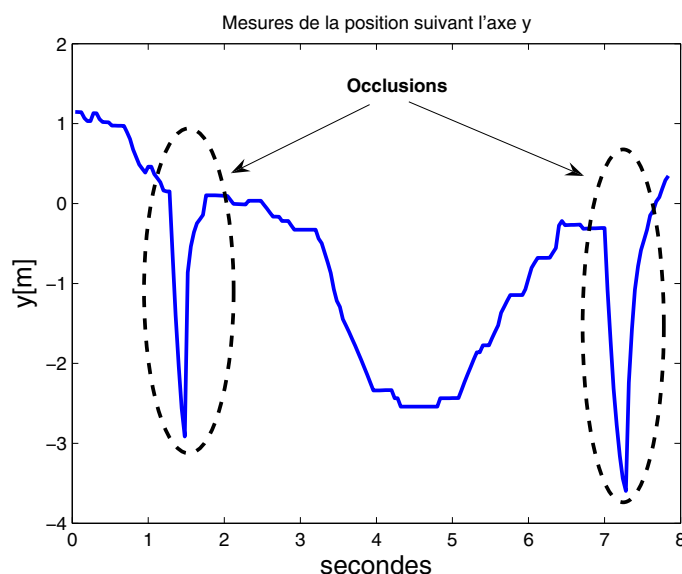


FIG. 2.21 – Influence des occultations sur la mesure de la position suivant  $y_h$ .

Cette contrainte sur le déplacement de la personne se traduit par l'inégalité suivante :

$$\begin{cases} |(x_k - x_{k-1})| \leq \alpha M_x \\ |(y_k - y_{k-1})| \leq \alpha M_y \end{cases} \quad (2.23)$$

avec  $M_x$  et  $M_y$  respectivement les moyennes de déplacement suivant les axes  $x_h$  et  $y_h$ , et  $\alpha$  un facteur de réglage du seuil suivant les conditions d'expérimentation. Nous choisissons  $\alpha = 4$  en se basant sur la vitesse moyenne d'un humain lorsqu'il marche afin de tolérer des déplacements rapides tout en rejetant les déplacements aberrants.

Les conditions expérimentales sont les mêmes que pour les simulations sans contrainte. Nous avons choisi un horizon d'estimation  $N_e = 5$ .

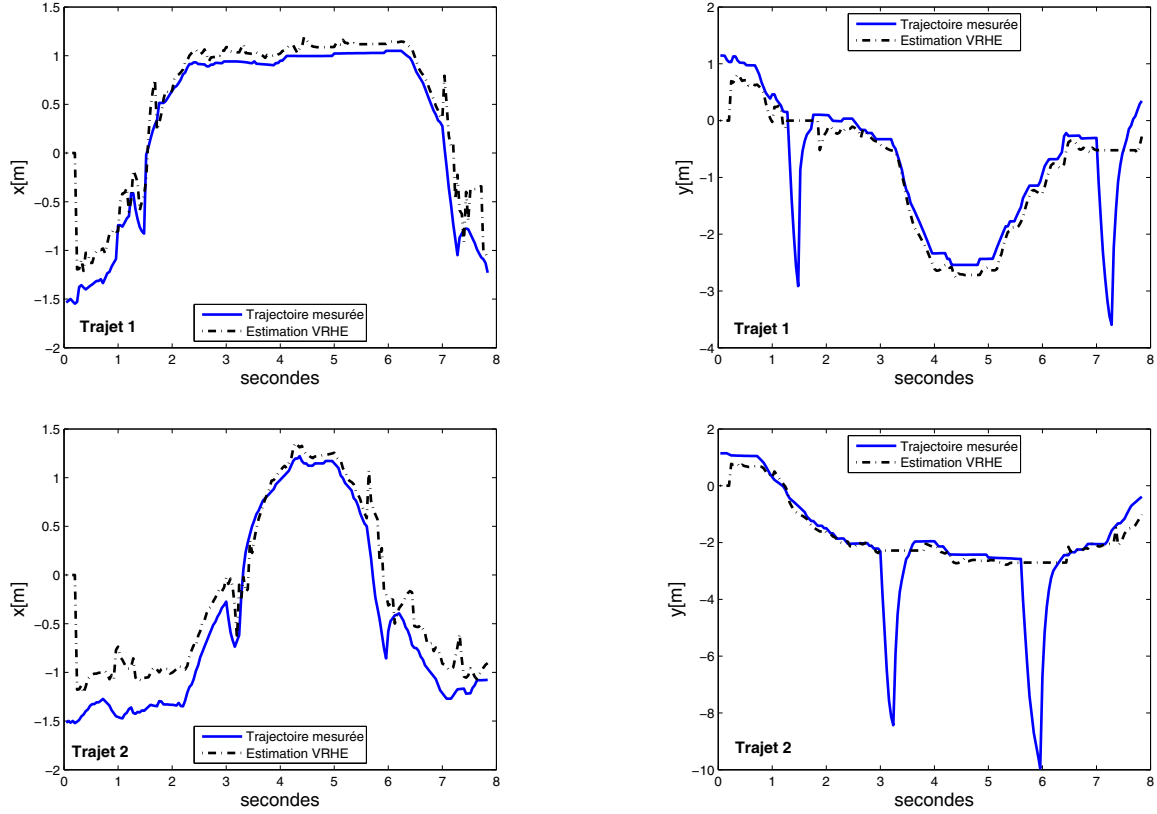


FIG. 2.22 – Exemples de prise en compte de contrainte sur le déplacement sur deux trajets.

La figure 2.22 illustre la prise en compte d'un déplacement admissible sur deux exemples de trajet. On constate qu'après l'occultation, grâce à la prise en compte de contrainte, le VRHE converge vers la position mesurée.

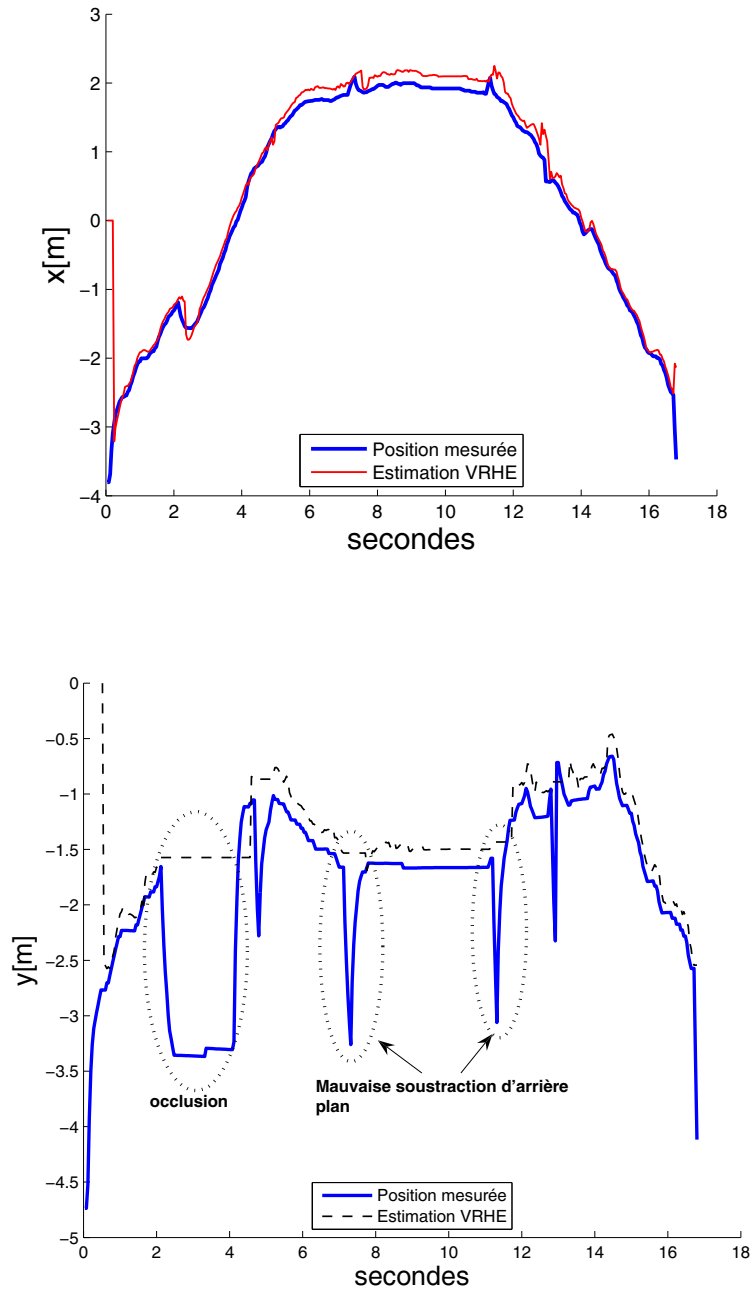


FIG. 2.23 – Prise en compte des variations lumineuses.

On remarque sur la figure 2.23 que cette contrainte de déplacement admissible permet également de gérer les changements de dimension brusques de la boîte englobante due à une mauvaise détection lors de la phase de soustraction d'arrière plan. Les variations lumineuses constituent l'une des principales perturbations lors de la phase de soustraction.

**Remarque 2.22** *La gestion des occultations peut aussi être envisagée en introduisant des*

*contraintes de visibilité relatives à la présence d'un obstacle dans l'image. Par contre, elle nécessite une extraction supplémentaire d'information à partir du traitement d'images. Le temps de calcul est actuellement trop long pour une approche temps réel.*

## 2.4 Conclusion

Dans ce chapitre nous avons essayé de montrer que l'estimation visuelle sur horizon glissant apporte une solution à la localisation 3D d'une personne à partir d'informations visuelles.

Le problème de la localisation 3D est reformulé en un problème d'optimisation non linéaire dans le plan image. A ce problème peuvent s'ajouter des contraintes portant sur la personne (limitation du déplacement admissible) ou sur l'image (contrainte de visibilité). L'estimation de la position 3D est assurée par un modèle global. Ce modèle combine le modèle du dispositif visuel et le modèle de déplacement humain. Dans notre cas, il s'agit d'un modèle de caméra perspective et d'un modèle de déplacement non holonome d'une personne. L'approche peut être utilisée quelque soit le système de vision ou de déplacement, lui conférant ainsi flexibilité et adaptabilité. Nous avons mis en évidence l'influence de l'horizon d'estimation sur la vitesse de convergence et le temps de calcul. Lors des expérimentations, la faisabilité et les potentialités du VRHE pour la localisation 3D ont été confirmées en présence d'un obstacle statique et en présence de bruits et de perturbations.





## Chapitre 3

# La fusion de données multi-capteurs

Détecter la présence, localiser et déterminer la posture d'une personne ne peut se faire avec un seul capteur. Pour le projet CAPTHOM, les technologies retenues sont l'infrarouge (détecteur IRP et thermopiles) et la caméra vidéo (caméra perspective fixe). Les informations issues de ces différents capteurs, et donc de nature hétérogène, doivent être traitées simultanément.

Nous allons montrer l'intérêt et la nécessité de fusionner ces données aussi bien pour renseigner les divers registres (détection de présence, localisation 3D, détermination de la posture) que pour prendre les décisions finales (alerte, déclenchement de l'éclairage ou du chauffage). Nous allons proposer, en fonction des spécificités de chaque objectif, une méthode de fusion.

Après avoir justifié la fusion de données pour le projet CAPTHOM, nous présentons les différentes architectures de fusions et les méthodes associées. Ensuite, nous présentons les solutions de fusion retenues pour chaque objectif en détaillant plus particulièrement la détermination de posture de la personne (sans capteur sur la personne).

Des résultats expérimentaux illustrent l'approche développée.

## 3.1 La fusion de données pour le projet CAPTHOM

Le développement d'un dispositif de fusion de données ne peut se faire sans une définition et une connaissance précise des objectifs de l'application car ils conditionnent en partie les choix en matière de méthodes mais également d'architecture globale. Il ne faut pas perdre de vue ces objectifs afin que le système de fusion délivre des informations pertinentes et utiles, facilitant ainsi la prise de décision. Il convient alors de se poser certaines questions avant de se lancer dans le développement d'une approche. La première d'entre-elles est de se demander quel est le but de la fusion de données pour l'application considérée. Avant de présenter le principe de la fusion de données et de décrire les principales méthodes et techniques existantes (§3.2), il nous semble important de préciser pourquoi la fusion de données est nécessaire pour le projet CAPTHOM et ce, quelle que soit l'application visée.

### 3.1.1 Pourquoi la fusion de données dans CAPTHOM ?

#### 3.1.1.1 L'obtention d'informations pertinentes

##### **La détection de présence humaine**

Problématique principale du projet CAPTHOM, la détection de présence humaine est une information nécessaire, commune à la gestion énergétique et au maintien à domicile. Comme nous l'avons décrite au Chapitre 1, la détection de présence est loin d'être triviale dans le sens où elle ne se limite pas à une simple détection de mouvement. Certes, le mouvement trahit la présence d'un individu mais l'absence de mouvement doit aussi être prise en compte. Le dispositif doit également être capable de discriminer les cibles détectées : seule la présence d'un humain doit être prise en compte. On a constaté qu'aucun capteur ou technologie ne permet de satisfaire les exigences de la détection de présence humaine de manière très fiable. Plusieurs capteurs sont donc nécessaires afin de délivrer différentes informations qui, fusionnées, permettent de statuer sur la présence ou non d'une personne, tout en minimisant l'impact des perturbations.

##### **La localisation 2D et 3D d'une personne**

Une autre demande du dispositif CAPTHOM est la localisation 2D et 3D de la personne. Nous avons montré dans le Chapitre 2 comment cette localisation en 3D pouvait être effectuée à partir de données de localisation 2D issues d'un algorithme de traitement de l'image. Là encore, on ne peut pas se fier uniquement aux données de la caméra pour plusieurs raisons. Tout d'abord, une personne immobile n'est plus détectée au bout d'un certain temps dans la mesure où elle est considérée comme appartenant à l'arrière plan.

La phase de soustraction d'images ne donne alors plus d'information. Seule l'analyse de données issues d'autres capteurs permet de conforter la détection de présence et la localisation de la personne. Cette analyse permet également de diminuer les fausses détections, comme le reflet d'une personne.

Par ailleurs, pour des raisons de performances énergétiques, il est préférable de coupler le dispositif visuel avec un dispositif de détection de mouvement afin d'éviter que la caméra ne fonctionne en permanence. Un détecteur IRP peut par exemple activer la caméra si il estime qu'un mouvement s'est produit. La caméra viendra alors confirmer ou infirmer l'hypothèse du détecteur IRP.

Une localisation moins précise peut être réalisée à l'aide d'un réseau de capteurs disposés dans l'habitat en vue de maximiser les zones de recouvrement. Encore une fois, la combinaison et la fusion des données provenant de plusieurs sources sont nécessaires.

### 3.1.1.2 La prise de décision

La prise de décision, quelle que soit l'application (gestion énergétique ou maintien à domicile), n'est réalisée qu'en tenant compte de plusieurs informations.

On distingue deux types d'informations : des informations communes et des informations spécifiques à chaque application. La figure 3.1 permet d'illustrer à la fois ce noyau commun mais également l'information spécifique aux deux approches.

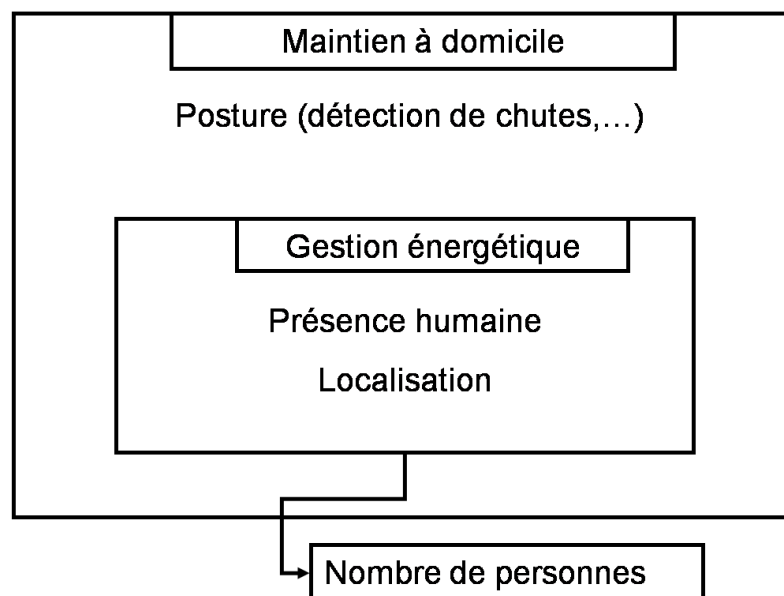


FIG. 3.1 – Objectifs du maintien à domicile et de la gestion énergétique.

### **Les informations communes :**

La détection de présence et la localisation de la personne sont nécessaires pour les deux applications de CAPTHOM. Des informations relatives à l'environnement extérieur telles que la météo ou l'horloge peuvent se révéler aussi fort utiles.

#### **• L'horloge**

L'horloge, ou plus précisément l'heure de la journée, est une information importante. En effet, certaines situations n'auront pas la même signification ou le même impact sur la décision finale en fonction de l'heure de la journée à laquelle elles sont apparues ou de la durée pendant laquelle elles ont eu lieu. Prenons par exemple le cas d'une personne âgée se rendant aux toilettes. Cette action, tout à fait banale pour la plupart des personnes, peut être source d'accident pour une personne âgée notamment lorsqu'elle s'assoit ou qu'elle se relève. Après avoir détecté la présence de la personne dans les toilettes, la prise en compte de la durée de la présence, grâce à une horloge, permet, si le séjour est trop long, de déclencher une alerte signifiant que quelque chose est probablement arrivé, notamment une chute.

#### **• La météo**

La météo constitue la seconde information issue de l'environnement extérieur, information que l'on retrouve dans plusieurs projets d'habitat intelligent. Cette information se révèle toutefois beaucoup plus pertinente dans une optique de gestion énergétique que de maintien à domicile. En fonction de la luminosité et/ou des conditions météorologiques, certaines opérations sont ou non effectuées. En effet, si la luminosité est suffisante, il apparaît inutile d'éclairer même si la présence d'une personne est détectée. A l'inverse, dès que la luminosité extérieure ne sera plus suffisante, il conviendra d'éclairer la pièce afin de continuer d'assurer le confort visuel de la personne si toutefois elle est encore présente dans la pièce.

***Remarque 3.1*** Ces deux informations symbolisent bien la définition assez large du terme « données » en fusion et montrent que tout phénomène doit être interprété en fonction de son contexte.

### **Les informations spécifiques :**

#### **• La posture**

La détermination de la posture de la personne (position debout, couchée, assise,...) est une information fondamentale pour la surveillance de personnes âgées et offre des possibilités supplémentaires de reconnaissance de situations de chute. Ainsi, plus la chute sera

rapidement détectée, plus court sera le temps d'intervention des équipes médicales. Le temps est un facteur déterminant sur ce type d'intervention. Parmi les accidents de la vie courante, la chute à domicile constitue la première cause de décès tout âge confondu [Erm07]. Les personnes âgées de 75 ans et plus sont encore plus concernées.

- **Le nombre de personnes**

La détermination du nombre de personnes présentes dans une pièce est une donnée essentielle pour la gestion énergétique optimisée de bâtiment. Cette donnée se révèle moins indispensable pour une application de maintien à domicile dans la mesure où la personne vit généralement seule dans son habitat. La présence éventuelle d'une autre personne, comme un personnel médical ou un parent, engendre une vigilance moindre du dispositif d'alerte, la personne prenant temporairement le relais.

Le tableau 3.1 résume l'ensemble des registres du dispositif CAPTHOM à renseigner, afin d'aider à la prise de décision en fonction de l'application. On retrouve les informations communes et les informations spécifiques.

Maintien à domicile	Gestion de l'énergie/Domotique
Présence humaine (personne en mouvement et immobile)	
Localisation 2D et/ou 3D	
Horloge (heure, durée)	
Météo (température, luminosité)	
Posture (debout, assis, accroupi, couché)	Nombre de personnes

TAB. 3.1 – Liste des registres à renseigner suivant l'application choisie.

Au regard de la variété des informations à renseigner, aucun capteur ou aucune technologie ne permet de renseigner l'ensemble de ces registres. La détection de présence humaine seule nécessite la combinaison de données de natures variées provenant de plusieurs détecteurs. Le besoin de développement d'une architecture de fusion se révèle donc nécessaire et sans mentionner comment la fusion sera effectuée, on peut d'ores et déjà dire qu'elle interviendra à plusieurs niveaux. Avant cela, nous allons répertorier les différents capteurs et les différentes technologies que nous allons utiliser pour renseigner ces registres.

### 3.1.2 Les données à fusionner

On distingue deux types d'informations :

- soit il existe un capteur qui fournit directement l'information. Dans ce cas, un simple traitement des données est suffisant ;
- soit plusieurs capteurs sont nécessaires. Une combinaison de leurs données doit être effectuée pour obtenir l'information utile.

On a vu dans le Chapitre 1 que la détection de la présence humaine ne pouvait pas être garantie par un seul capteur ou une seule technologie. Au regard des limitations technologiques, des spécifications du cahier des charges (moindre coût, couverture spatiale et temporelle) et des contraintes liées au facteur humain (dispositif non invasif et non intrusif, respect de la vie privée), les détecteurs infrarouges passifs, les thermopiles et les caméras constituent les trois technologies retenues pour répondre aux exigences du dispositif CAPTHOM en terme de détection de présence. La figure 3.2 permet d'illustrer l'apport de chacune de ces technologies à la fois pour des applications de maintien à domicile et de gestion énergétique, soulignant une nouvelle fois le caractère commun de ces deux approches et leurs spécificités.

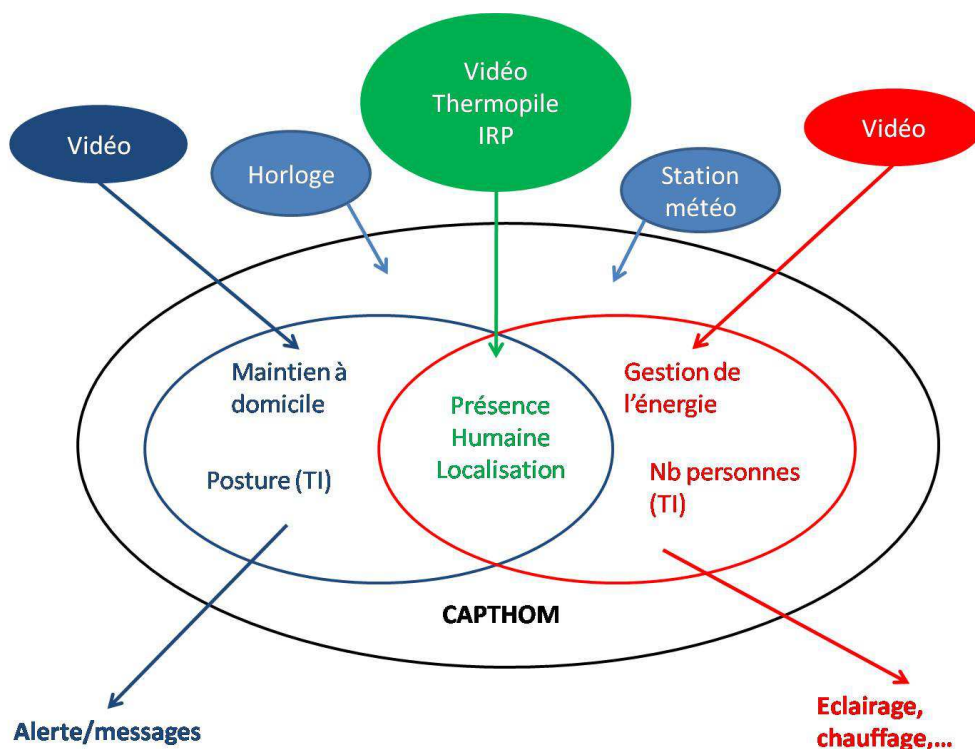


FIG. 3.2 – Le dispositif CAPTHOM.

Le fonctionnement du dispositif CAPTHOM repose sur un système multi-capteurs de technologies différentes : chacune apportant son lot de données à traiter et à partir desquelles des décisions doivent être prises. L'exploitation de ces différentes données nécessite le développement d'une architecture de fusion, facilitant le travail d'extraction des données pertinentes, diminuant les bruits de mesures et gérant au mieux les imperfections accompagnant généralement tout type de données.

On s'intéresse maintenant au concept de fusion de données et on présente un bref tour d'horizon des différentes méthodes en soulignant leurs avantages et leurs inconvénients. Ceci permettra de justifier le choix des méthodes retenues pour l'expérimentation.

## 3.2 Généralités en fusion de données

Concept émergeant dans les années 1970 et à l'origine pour des applications militaires, la fusion de données fait depuis, l'objet d'une attention particulière. Ses domaines d'applications sont très variés : maintenance de procédé, diagnostic médical, surveillance d'environnement, robotique mobile, ... L'observation de certains phénomènes nécessitent de plus en plus de combiner des informations variées provenant d'un réseau hétérogène de capteurs au lieu de considérer l'utilisation de ces derniers de manière séparée, en étudiant les données provenant d'une source indépendamment des autres sources. La combinaison de ces données permet, en général, d'obtenir une information beaucoup plus riche et plus précise du phénomène observé à un instant  $t$ . Les différentes sources sont même susceptibles d'apporter des informations supplémentaires afin de confirmer une hypothèse ou au contraire de diminuer sa probabilité voire de l'écarter définitivement. Différentes approches sont envisageables. On peut fusionner soit des informations différentes à l'instant  $t$ , soit une même information à des instants différents. On peut combiner des données provenant de différents capteurs mais également des connaissances a priori provenant de bases de données, de résultats antérieurs de fusion, d'avis d'expert, ...

***Remarque 3.2*** *Si on fait le parallèle avec le comportement humain (ou animal), on est souvent confronté à des problèmes de fusion de données. En effet, nous disposons d'un réseau de « capteurs » qui nous permettent, par la combinaison de leurs informations, de prendre telle ou telle décision face à une certaine situation. A travers la vue, le toucher ou encore l'ouïe, on va par exemple pouvoir détecter une situation de danger et même si l'un de ces sens nous fait défaut, les autres nous permettent d'établir un diagnostic de la situation afin de prendre une décision. Nos réactions et nos décisions dépendent également de notre expérience, de notre savoir.*



Dans un premier temps, nous allons définir les notions de « fusion » et de « données » que nous considérerons par la suite (leurs définitions étant parfois sujettes à interprétation). Les apports de la fusion et les caractéristiques des données sont décrits afin de bien cerner les avantages et les difficultés rencontrés en fusion de données multi-capteurs. Dans une seconde partie, la modélisation et le principe général de fonctionnement de la fusion sont présentés, introduisant ainsi les notions de niveau de fusion et d'architecture. Enfin, un bref panorama des méthodes et théories utilisées en fusion, des approches numériques aux approches symboliques, est présenté.

### 3.2.1 Définition de la fusion de données

La fusion de données a longtemps été pénalisée par la diversité de ses applications et par l'apparition, très tardive, d'une définition suffisamment générale. En effet, des notions et des termes fondamentaux tels que « fusion de données » voire même le terme « donnée », possédaient des définitions différentes d'un domaine à l'autre. Cette absence de formalisme et d'uniformité terminologique a longtemps été préjudiciable à la fusion de données.

Afin de corriger cette situation, le JDL, pour Joint Directors of Laboratories Data Fusion Group, a été créé en 1985 dans l'optique d'enclencher un processus de codification de la terminologie autour de la fusion de données et de simplifier ainsi les discussions et les transferts de technologies entre les différents domaines. La première définition donnée par le JDL, en 1985 était la suivante [Lig09] :

*« Processus traitant de l'association, la corrélation et la combinaison des données et de l'information provenant d'une source unique ou de plusieurs sources afin d'obtenir des estimées affinées de positions et d'identités et l'évaluation complète et au bon moment des situations et des menaces, ainsi que leur importance. Le processus est caractérisé par une amélioration continue des estimées et des appréciations, et par l'évaluation du besoin en sources additionnelles, ou de modification du processus lui-même, afin d'obtenir des résultats améliorés. »*

Cette définition a bien entendu évolué depuis sa création et a fait l'objet de critiques notamment concernant sa longueur mais également sur le côté restrictif du terme « évaluation des menaces » qui doit être étendu à l'impact de n'importe quelle situation estimée. Malgré tout, cette définition permet d'illustrer le principe de la fusion en termes d'organisation, de combinaison et d'interprétation des données et des informations provenant de

plusieurs capteurs. La fusion de données multi-capteurs apporte des avantages indéniables que nous décrivons ci-dessous :

- **Une amélioration de l'estimation de l'état d'une cible.** Dans le cas du suivi d'une personne dans une maison à partir de données provenant uniquement de détecteurs infrarouges passifs, l'utilisation de plusieurs capteurs intelligemment disposés dans l'environnement permet une estimation plus fine de la position de cette personne. Dans le même ordre d'idée, la combinaison des observations successives d'un même capteur durant une certaine fenêtre temporelle conduit à des estimées plus précises de l'état observé ;

***Remarque 3.3** L'estimation visuelle sur horizon glissant peut être vue comme une méthode de fusion de données dans la mesure où elle améliore la connaissance de l'état de la personne (estimation du vecteur d'état) à partir d'un nombre fini de mesures passées.*

- **Une amélioration du dispositif d'observations.** Plus la quantité de données disponibles est importante, plus la prise de décision est facilitée. On a donc tout intérêt à utiliser le plus de capteurs possibles en tenant compte, tout de même, des contraintes liées au placement optimisé, au coût, à la fiabilité ;
- **Une amélioration des observations.** En diversifiant le type de capteurs utilisés pour l'observation d'une même situation, on élargit d'autant le spectre des observations physiques, créant alors des relations de complémentarité ou de redondance. On considère un dispositif combinant un détecteur infrarouge passif et un détecteur ultrason, combinaison relativement classique dans des dispositifs commercialisés. Le détecteur IRP peut être perturbé par des sources de chaleur rendant la détection de mouvement plus difficile. Le détecteur ultrason ne sera pas affecté par ces perturbations, mais rencontrera des difficultés lors de mouvements transversaux, situation moins problématique pour le détecteur infrarouge. L'association de ces deux technologies permet d'améliorer la fiabilité de la détection et de diminuer l'impact des perturbations.

## Modification de la définition du JDL

La fusion de données s'est étendue à d'autres domaines d'applications [Col01][Bel02] que celui militaire, rendant alors la définition originelle un peu trop restrictive aussi bien en termes de possibilités d'utilisation que de vocabulaire. Des révisions à cette définition ont donc été apportées [Ste09] parmi lesquelles nous pouvons retenir les suivantes :

1. Le concept de *combinaison de données* englobe un large éventail de problèmes, la corrélation n'en fait pas partie. En effet, la corrélation, en probabilité ou en statistique, est une des méthodes permettant d'évaluer le lien hypothétique entre des données.
2. La référence à *l'estimation de position et d'identité* doit être étendue afin de couvrir l'ensemble des variétés d'estimation d'état.
3. *L'évaluation des menaces* restreint les applications à des situations où la menace est un facteur alors que la définition doit être étendue afin de prendre en compte l'impact que peut avoir n'importe quelle situation estimée.
4. Tout processus de fusion n'implique pas forcément la mise en place d'un dispositif de gestion de collecte d'informations ni d'ajustement du processus.

Suite à ces révisions par rapport à la définition initiale du JDL, [Ste09] propose donc la définition simplifiée suivante :

**Définition : 3.1** *Processus combinant données et informations afin d'estimer et de prédire les états d'une entité.*

### 3.2.2 Les données, élément essentiel de la fusion

D'après la définition du JDL, la fusion permet de gérer des informations hétérogènes, hétérogénéité qui peut se traduire de différentes manières.

#### 3.2.2.1 Les caractéristiques des données

Une première différence entre les données se fait au niveau de la forme de ces dernières. Le terme « données » doit être vu dans le sens le plus large possible et pas uniquement comme l'ensemble des mesures provenant de capteurs. En effet, l'utilisation de base de données, de résultats de fusion antérieurs, d'avis d'expert ou de connaissances génériques exprimées sous forme de règles doit être prise en compte et englobée dans le terme « données ». Les données peuvent être numériques (données brutes ou traitées provenant généralement de capteur) ou symboliques (sous forme de mots ou de phrases). Cette hétérogénéité au niveau de la forme oblige à rechercher des mécanismes pour ramener l'ensemble des informations dans un référentiel commun, indispensable à leurs combinaisons.

**Remarque 3.4** *Il n'est cependant pas toujours possible de ramener l'ensemble des données dans un même référentiel voire pas toujours utile. En effet, on distingue différents niveaux de fusion qui vont dépendre du type de données à combiner. On parle de fusion*

*« bas niveau » dès lors que l'on cherche à combiner des données brutes, directement issues des capteurs et de fusion « haut niveau » quand on travaille sur des données ayant subies des étapes préalables de traitement. Plus on se dirige vers un niveau élevé d'abstraction, plus le niveau de fusion augmente et implique des méthodes et des techniques différentes.*

Le panel de données dont on dispose pour l'évaluation et la caractérisation d'une situation a une influence sur l'architecture et l'enchaînement des différentes étapes de fusion afin d'utiliser, au mieux, le potentiel informatif de chaque type de données quel que soit son niveau et sa forme.

Toute donnée est entachée d'imperfections. Cette notion d'imperfection des données est inhérente au concept de fusion de données car elle donne tout son sens au besoin de combiner des informations afin de diminuer l'impact de leurs imperfections. La fusion n'a de sens que si les données que l'on manipule sont entachées de défaut, car si les données étaient parfaites les combiner n'aurait aucune utilité.

### 3.2.2.2 Les imperfections des données

L'un des intérêts de la fusion de données est de minimiser l'impact de l'imperfection des données. Toute donnée est entachée d'imprécision ou d'incertitude voire des deux.

- **L'incertitude** est relative à la véracité d'une information et caractérise son degré de conformité à la réalité. Référence à la nature de l'objet ou du fait observé, l'incertitude porte davantage sur la qualité de l'information que sur son contenu. Même si une donnée est précise, sa réalisation n'est pas forcément sûre : la fiabilité des capteurs utilisés ayant également un impact sur l'incertitude.

- **L'imprécision** porte sur le contenu de l'information et correspond à un défaut quantitatif et non qualitatif. Elle peut se définir comme un manque d'exactitude en quantité mais également comme une difficulté d'interprétation d'une proposition mal définie. A partir du moment où on est confronté à une insuffisance de connaissances sur la donnée, on parle d'imprécision. Sa description passe alors par l'utilisation de termes flous.

**Remarque 3.5** *On a souvent tendance à confondre ces deux notions et de qualifier d'imprécision ce qui correspond en réalité à de l'incertitude. Il existe une sorte de dualité entre incertitude et imprécision puisque l'on ne peut privilégier l'une qu'au détriment de l'autre. Soit l'exemple d'une grandeur évaluée par un intervalle. Plus on réduit l'intervalle, entraînant une augmentation de la précision, moins on aura de chance que la vraie valeur*

*soit comprise dans cet intervalle. La proposition sera alors de moins en moins certaine. A l'inverse, si on augmente la taille de l'intervalle, on va gagner en certitude ce que l'on va perdre en précision.*

- **L'incomplétude** : cette imperfection caractérise l'absence d'informations apportée par la source sur certains aspects du problème. D'ailleurs, c'est l'incomplétude des informations issues de chaque source qui motive la mise en place d'un processus de fusion. L'information fournie par une source n'est en général que partielle, ne fournissant qu'une vision spécifique de l'environnement ou du phénomène observé. L'information délivrée par un détecteur IRP permet de caractériser le mouvement mais ne fournit aucune indication concernant l'identité de la cible détectée par exemple.

- **L'ambiguïté** : elle caractérise la capacité d'une information à conduire à deux interprétations différentes. Elle peut être la résultante d'autres imperfections. L'imprécision d'une mesure peut entraîner des difficultés à différencier deux situations. De même, une information incomplète peut induire une confusion entre des objets qui ne peuvent pas être distingués au vue des caractéristiques mises en évidence par la source. L'un des objectifs de la fusion est notamment de lever ces ambiguïtés en s'appuyant sur les informations provenant d'autres sources.

- **Le conflit** : cette notion caractérise le fait que plusieurs informations conduisent à des interprétations contradictoires et donc incompatibles. Les situations de conflit sont fréquentes en fusion de données et aboutissent généralement à des problèmes dont la résolution est vite contraignante. Pour lutter contre ces situations de conflit, on peut, soit éliminer les sources non fiables, soit retarder la combinaison des informations en attendant une nouvelle donnée susceptible de lever le conflit.

On voit donc clairement l'intérêt de la fusion car pour évaluer des situations complexes, on ne peut faire confiance à une seule source d'informations, mais plusieurs données doivent prises en compte.

Heureusement les sources d'informations ne présentent pas que des imperfections mais possèdent également d'autres caractéristiques ayant un impact positif sur le processus de fusion :

- **La redondance** : la redondance est la qualité de sources qui apportent chacune le même type d'informations. Les informations redondantes permettent alors de diminuer l'imprécision ou l'incertitude des données.

- **La complémentarité** : c'est, par définition, la propriété des sources qui apportent des informations sur des grandeurs différentes d'un même objet. La mise en commun de ces informations permet d'obtenir une information plus complète et ainsi lever les éventuels problèmes d'ambiguïté que l'on peut rencontrer.

**Remarque 3.6** *L'association d'une thermopile et d'un dispositif visuel, dans l'application CAPTHOM, permet d'obtenir une redondance concernant la présence d'une entité. Le dispositif visuel vient compléter l'information de présence en caractérisant l'entité détectée.*

### 3.2.3 Modélisation du processus de fusion de données

Dans leur optique de clarifier et codifier le langage autour de la fusion de données, le Joint Directors of Laboratories (JDL) Data Fusion Working Group s'est également occupé de modéliser le processus de fusion de données de telle manière qu'il soit suffisamment général afin d'être utilisable par plusieurs domaines d'application. Le modèle proposé par le JDL est décrit dans [Hal09][Kes09] et une révision de ce modèle est donnée dans [Ste09].

Le modèle présenté dans [Hal09], illustré par la figure 3.3, correspond au modèle original proposé par le JDL. La fusion de données est vue comme la mise en relation de sources avec une interface homme-machine, faisant appel à différents niveaux de fusion (5 dans le modèle original) et à un système de gestion de données.

#### 3.2.3.1 Les niveaux de fusion

Différentes classifications en niveaux ont été proposées, chacune reposant sur une approche différente selon le point de vue choisi. Une approche centrée sur le processus peut être envisagée avec un nivellement de type : identification, suivi, association, prédiction. On peut également considérer les niveaux en fonction du sujet : objets, situations, menaces. Enfin, on peut envisager un point de vue centré sur les produits : objet, situation, coût de l'estimation. [Ste09] propose une partition suffisamment claire et utile à l'ensemble de la communauté de la fusion de données. Les différents niveaux de fusion proposés sont en fait des relectures des niveaux proposés par le modèle du JDL avec, comme pour la définition du processus de fusion, certaines corrections. Ces niveaux sont les suivants :

- Niveau 0 : *Prétraitement des données*. Estimation du signal et d'états caractéristiques, afin d'atténuer le bruit, d'opérer des alignements spatiaux ou temporels. Exemple : réaliser l'extraction de vecteurs caractéristiques grâce au traitement d'images.

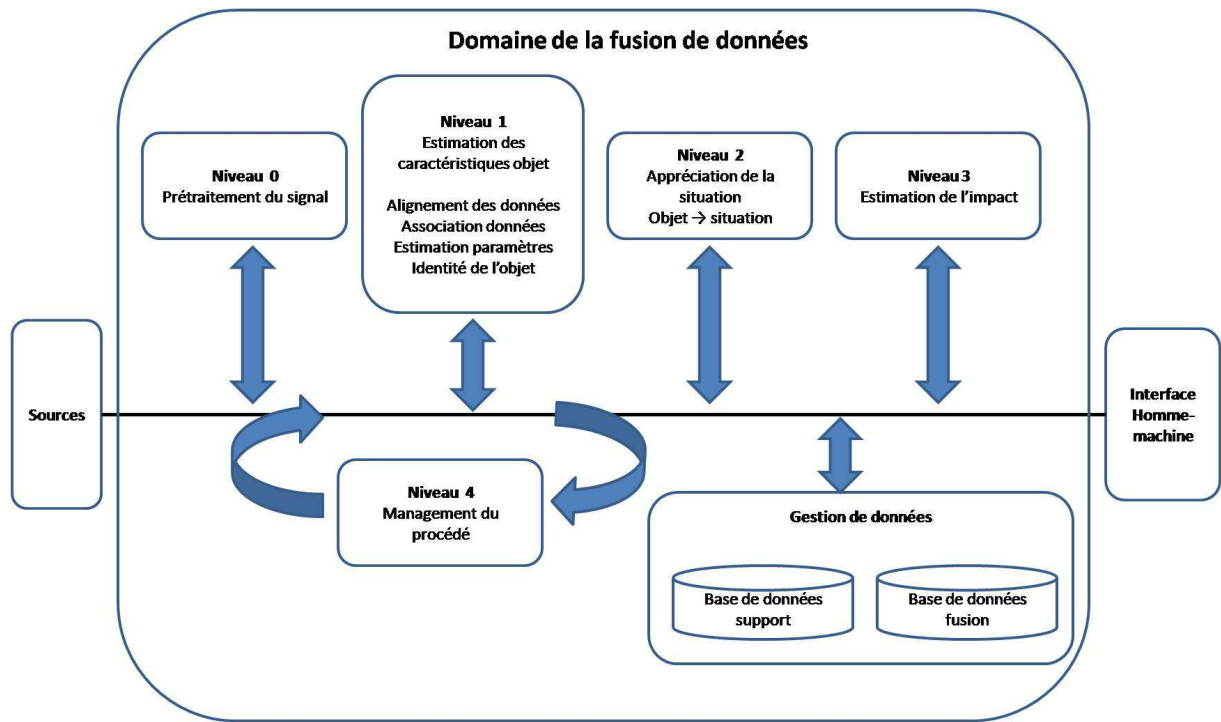


FIG. 3.3 – Modèle de la fusion de données défini par le JDL [Hal09].

- Niveau 1 : *Estimation de l'entité*. Association de données afin d'estimer les paramètres et les attributs de l'entité observée, comme sa position, sa vitesse ou encore son identité.
- Niveau 2 : *Appréciation de la situation*. Analyse et évaluation des relations entre les différentes entités présentes au sein de l'environnement d'observation et de leurs implications sur les états de l'entité observée.
- Niveau 3 : *Estimation de l'impact*. Analyse de la situation afin d'estimer et de prédire les effets et les conséquences.
- Niveau 4 : *Management du procédé*. Système d'auto-évaluation des performances du dispositif de fusion (éventuels besoins supplémentaires de contrôle ou ajout de nouvelles sources de données).

**Remarque 3.7** Un niveau 5 été introduit par [Hal00] et [Bla03]. Il vise à améliorer l'interaction entre le système de fusion et un utilisateur afin d'aider à l'analyse des résultats et à la prise de décision.

Le niveau 1 peut être subdivisé suivant 3 processus dont une représentation générale est proposée dans [Cro93] et illustrée à la figure 3.4. Après un alignement temporel et/ou spatial de l'ensemble des données et observations (niveau 0), une mise en correspondance (association des données) des données entre elles et avec les prédictions issues de la phase

précédente est effectuée. Chaque information doit être associée à la bonne entité ou au bon attribut d'une entité particulière. On procède ensuite à des estimations des états du modèle afin d'obtenir une mise à jour de la description de l'entité observée. Cette mise à jour offre la possibilité d'ajouter de nouvelles informations mais également de supprimer celles qui seraient devenues obsolètes, ou à l'origine d'ambiguïté ou de conflit. Enfin, on effectue une prédiction de l'identité ou de la classification de l'entité.

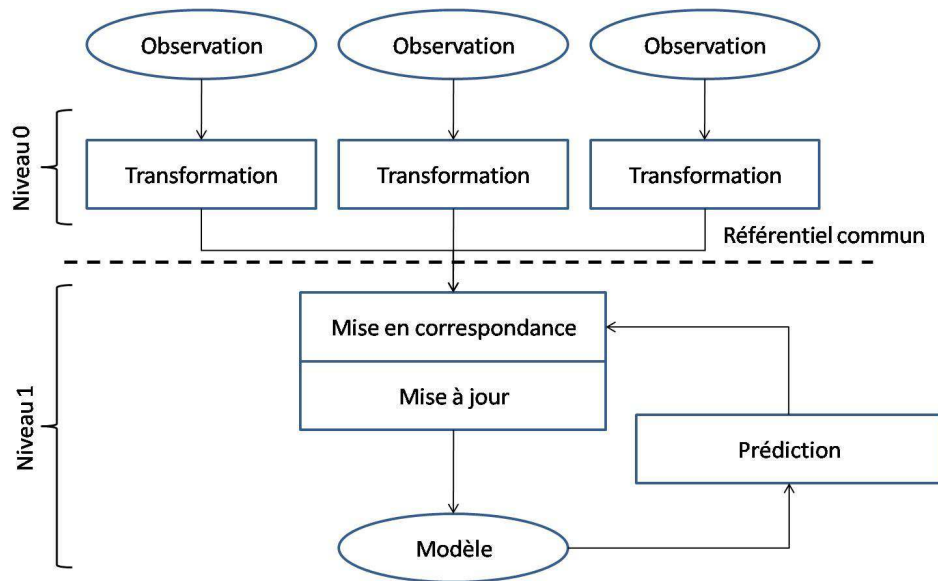


FIG. 3.4 – Représentation des processus de fusion de niveau 0 et 1 [Cro93].

Certaines erreurs d'appréciations peuvent être commises à la lecture du modèle proposé par le JDL. La première d'entre elles est de considérer la fusion de données comme un processus totalement indépendant. En théorie et dans la pratique, cette conception est totalement impossible. En effet, la fusion de données comme le précise [Kes09] n'est qu'un élément d'un ensemble de procédés tous nécessaires à la production d'informations utiles à un problème particulier ou à une prise de décision. [Kes09] décrit en effet un cycle que les auteurs nomment Information Processing Cycle (IPC), cycle qui intègre le processus de fusion avec d'autres procédés tels que le management des ressources, la fouille de données et l'environnement. Cette constatation rejoint en quelque sorte les révisions illustrées dans [Ste09] qui présente une modification du modèle du JDL issue des réflexions de [Bow04]. Le modèle, décrit figure 3.5, montre bien la prise en compte du management des ressources mais également le fait que la gestion des bases de données est extérieure au processus de fusion contrairement au modèle proposé par le JDL.



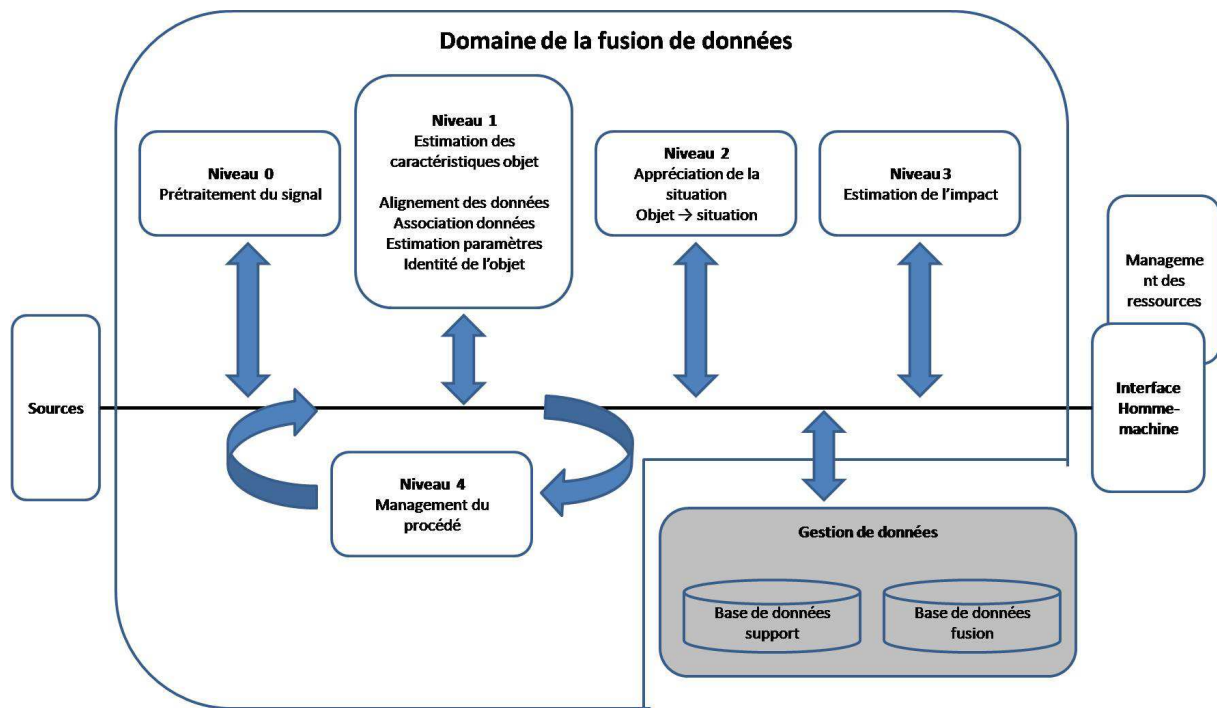


FIG. 3.5 – Révision du modèle du JDL [Bow04].

La seconde erreur d'interprétation qui peut ressortir du modèle du JDL est de ne considérer l'homme que comme une interface du processus de fusion. On cherche en effet de plus en plus à développer des systèmes fortement automatisés qui permettent d'effectuer la majorité des fonctions, minimisant ainsi le rôle de l'humain qui n'a alors plus qu'un rôle de lecture voire plus aucune implication. L'homme a cependant un rôle important à jouer dans un processus de fusion car son analyse repose sur son expérience et sur sa connaissance de l'environnement dans lequel se trouve l'entité observée. Reproduire les réflexions humaines de manière totalement automatisée n'est pas encore totalement acquis.

Enfin, le modèle de fusion tel qu'il a été défini par le JDL ne doit pas être vu comme un modèle imposant une hiérarchie et/ou une linéarité des processus. En ce sens, le terme *niveau* est peut être inapproprié comme le souligne [Kes09] car il accentue la vision de la fusion comme une séquence hiérarchisée de processus. Or, même si une certaine hiérarchie peut exister entre les différents niveaux, il n'est pas rare que des processus appartenant à différents niveaux soient exécutés en parallèle. De même, les résultats obtenus à un certain niveau de fusion ne sont pas forcément utilisés par le niveau situé au dessus. L'architecture d'un dispositif de fusion est donc composée à la fois de parties où les niveaux de fusion sont hiérarchisés ce qui est généralement le cas pour les niveaux 0 à 2 mais également de parties où un niveau agit de manière totalement indépendante.

### 3.2.3.2 Les architectures de fusion

Différentes architectures de fusion sont envisageables en fonction du niveau d'abstraction des informations. On distingue ainsi trois principaux types d'architecture de fusion [Lig09] :

1. La fusion de mesures : fusion directe des données capteurs (figure 3.6).
2. La fusion de primitives : représentation des données capteurs par des vecteurs caractéristiques qui sont alors fusionnés (figure 3.7).
3. La fusion de décisions : après calculs sur les données issues des capteurs, on aboutit à des inférences ou des décisions de plus haut niveau, décisions qui sont alors fusionnées (figure 3.8).

Le choix de l'architecture va naturellement dépendre du type de données dont on dispose mais également de l'objectif visé. Si des capteurs mesurent le même phénomène physique comme par exemple deux détecteurs IRP ou deux dispositifs de vision, on optera pour une fusion directe des données brutes de capteurs. Dès lors que les données provenant des capteurs sont dissemblables, la mise en place des deux autres types de fusion est inévitable.

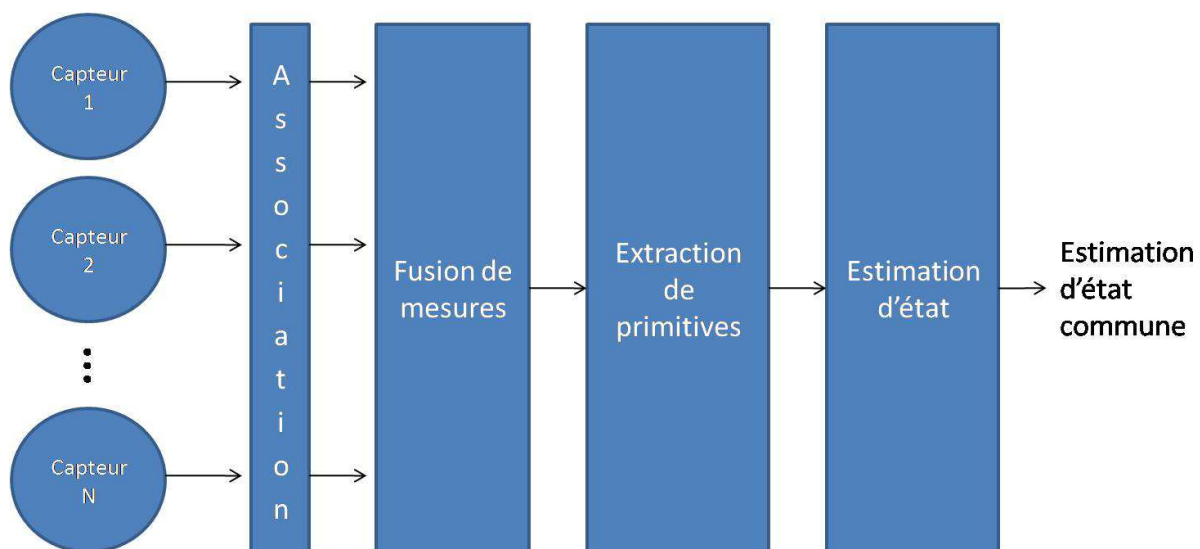


FIG. 3.6 – Représentation de la fusion de mesures.

Niveau de fusion, architecture, objectif sont donc autant de critères à prendre en compte lors du choix de la méthode ou de la technique de fusion. Plusieurs méthodes peuvent intervenir dans un même système à des moments différents.

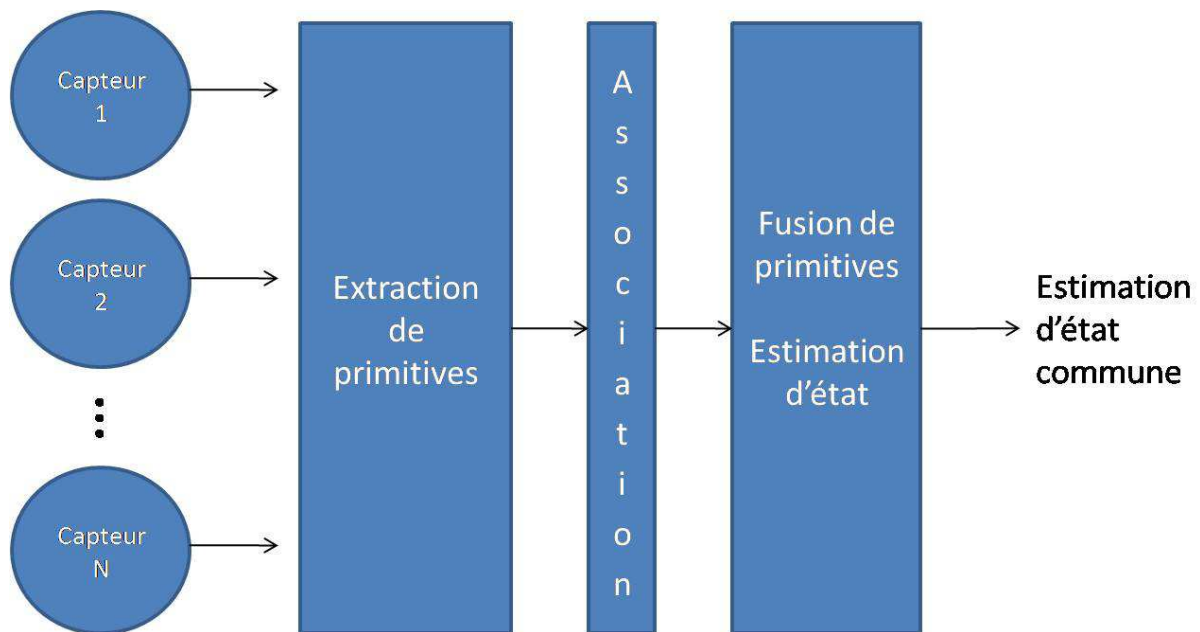


FIG. 3.7 – Représentation de la fusion de primitives.

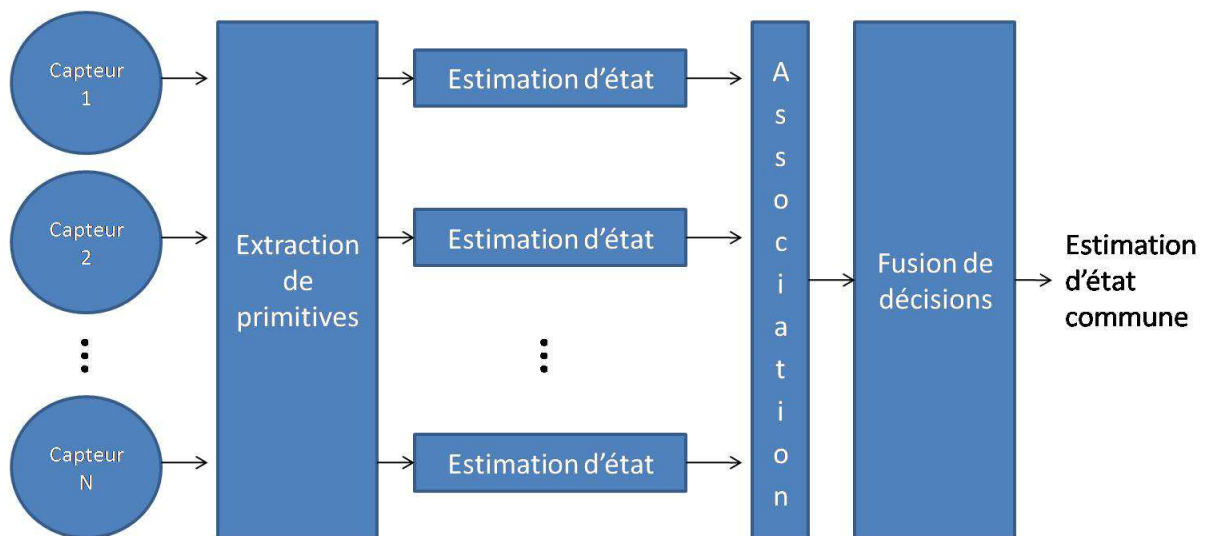


FIG. 3.8 – Représentation de la fusion de décisions.

### 3.2.4 Les théories et méthodes utilisées en fusion

Il est difficile d'associer une méthode à un type d'application ou à un niveau de fusion. On distingue deux grandes classes de méthodes : les méthodes probabilistes et numériques qui sont utilisées principalement aux niveaux 0 et 1, et d'un autre côté les méthodes symboliques, basées sur la manipulation de données plus textuelles que numériques, que l'on trouve essentiellement aux niveaux 2, 3 et 4.

### 3.2.4.1 Les approches probabilistes et numériques

#### Les approches probabilistes

L'approche par la théorie des probabilités est l'une des plus anciennes et celle qui est encore le plus souvent utilisée que ce soit en fusion de données ou en fusion d'images car elle repose sur des fondements mathématiques solides.

L'utilisation du théorème de Bayes permet de mettre à jour la probabilité d'occurrence d'une hypothèse en fonction de ses précédentes occurrences et des nouvelles observations. Cette technique peut utiliser des probabilités classiques ou subjectives.

Une distribution de probabilités décrit l'occurrence d'un événement passé, appelée *probabilité a priori*. A chaque pas d'échantillonnage, les observations modifient cette distribution de probabilités, qui devient alors une *probabilité a posteriori*. La formule de Bayes permet d'estimer les probabilités a posteriori des hypothèses envisagées ( $H_1$  à  $H_n$ ), hypothèses exhaustives et mutuellement exclusives, suite à l'arrivée d'une nouvelle observation.

Il s'agit ensuite de choisir le critère de décision. Le critère du maximum a posteriori est souvent choisi. L'hypothèse retenue est donc celle dont la probabilité a posteriori est la plus grande.

L'utilisation des probabilités permet une bonne représentation du caractère incertain des données, et est donc bien adaptée pour modéliser la fiabilité des capteurs. Par contre, les approches basées sur les probabilités n'assurent pas la gestion de l'imprécision.

**Remarque 3.8** *La génération d'alerte pour le maintien à domicile nécessite souvent la manipulation de termes flous afin de prendre en compte l'imprécision de certaines données. D'autre part, des connaissances a priori du système sont nécessaires et leurs expressions sous forme de probabilités n'est pas toujours possible. Enfin l'ignorance n'est pas du tout prise en compte. Elle se traduit ici par une égalité des probabilités a priori, ce qui entraîne parfois certaines incohérences.*

#### Les approches numériques

Appliqués à des données de bas niveau (mesures brutes des capteurs), les estimateurs (filtre de Kalman, filtre particulaire) permettent de réaliser l'estimation de l'état d'une cible. Il s'agit donc de méthodes de reconstruction de l'environnement ou de la dynamique des systèmes. Ces estimateurs, à partir de données multiples prises au même instant, permettent de résoudre des problèmes de localisation en fournissant une estimation de position ou de

vitesse. En revanche, pour des problèmes de suivi de cible, ces estimateurs s'appliquent à des séquences temporelles de données.

Parmi ces estimateurs, le filtre de Kalman est celui que l'on retrouve le plus souvent. À partir d'observations, il permet d'estimer la valeur d'un vecteur d'état en tenant compte des bruits de mesures et des bruits inhérents au système. L'estimation optimale du vecteur d'état, au sens des moindres carrés, est obtenue par itérations. En fusion de données, les mesures des différents capteurs sont considérées comme autant d'observations du même vecteur d'état. Si les résultats théoriques obtenus sont généralement satisfaisants, certaines hypothèses s'avèrent en pratique difficilement vérifiables (bruits blanc gaussiens non corrélés). Enfin on constate que les approches basées sur les estimateurs utilisent généralement des données provenant de capteurs de même nature.

**Remarque 3.9** *Pour des applications de type gestion de l'énergie ou maintien à domicile, on est confronté à une grande diversité de capteurs et de données. L'utilisation de ces méthodes est donc restreinte à des problématiques de localisation et de suivi de cibles à partir de données issues d'un réseau de capteurs identiques.*

Lorsqu'il s'avère nécessaire de travailler avec des données plus abstraites, des approches moins mathématiques telles que les méthodes symboliques, permettent de modéliser à la fois l'incertitude et l'imprécision.

### 3.2.4.2 Les approches symboliques

Les approches symboliques sont appliquées sur des données de plus haut niveau : extraction de primitives ou de caractéristiques, informations issues d'avis d'experts ou de retour d'expérience.

#### **Théorie de l'évidence**

La théorie de l'évidence a été introduite par Dempster [Dem67] [Dem68] puis reprise ensuite par Shafer [Sha76] lui conférant alors un formalisme mathématique plus abouti. Cette théorie est connue sous le nom de théorie de Dempster-Shafer ou de théorie des croyances. C'est une généralisation de l'inférence bayésienne avec un traitement de l'incertain, dans le sens où elle permet de représenter l'incertitude en considérant des événements qui ne sont pas obligatoirement exclusifs ni exhaustifs.

La théorie des croyances nécessite la définition d'un ensemble non nécessairement exhaustif d'hypothèses  $H_i$ , que l'on nomme *cadre de discernement* noté  $\Omega$ . L'objectif du problème consiste à identifier une situation  $x$  parmi ce cadre de discernement. Il est parfois

difficile de déterminer la liste exhaustive des hypothèses relatives à un problème donné, on parle alors de « monde ouvert ». Du cadre de discernement  $\Omega$ , on en déduit l'ensemble noté  $2^\Omega$  comprenant l'ensemble des  $2^N$  sous-ensembles  $A$  de  $\Omega$  défini de la manière suivante :

$$2^\Omega = \{A/A \subseteq \Omega\} = \{\emptyset, H_1, \dots, H_n, H_1 \cup H_2, \Omega\}$$

On assigne ensuite à chaque sous-ensemble  $A$  une masse de croyance  $m(A)$  traduisant le degré de confiance que l'on a pour chaque sous-ensemble  $A$ . Cette fonction  $m$ , appelée distribution de la masse de croyance, est définie par :

$$\begin{aligned} m : 2^\Omega &\rightarrow [0, 1] \\ A &\mapsto m(A) \end{aligned} \tag{3.1}$$

avec  $\sum_{A \in 2^\Omega} m(A) = 1$ .

Pour un problème donné, il s'agit ensuite de modéliser la connaissance de chaque source d'informations avec les différentes grandeurs définies sur  $2^\Omega$ , en associant une masse de croyance pour chaque sous-ensemble  $A$ . L'objectif est d'obtenir une distribution de masse de croyance finale qui prend en compte l'ensemble des informations, en fusionnant l'ensemble des masses de croyances de chaque source. Trois critères de décision peuvent être envisagés pour déterminer la classification :

1. La distribution de masse de croyance  $Crit(A) = m_{fusion}(A)$
2. La fonction de crédibilité  $Crit(A) = Cr(A) = \sum_{B \subseteq A} m_{fusion}(B)$
3. La fonction de plausibilité  $Crit(A) = Pl(A) = \sum_{A \cap B \neq \emptyset} m_{fusion}(B)$

L'avantage de cette méthode réside dans sa grande souplesse d'adaptation au cas par cas. En revanche, il n'existe pas de méthode générique de modélisation des connaissances. Il est nécessaire de s'appuyer sur des données expérimentales et des avis d'experts pour parvenir à modéliser les connaissances sous forme de distribution de masses. La réalisation de campagnes de tests permet d'effectuer les réglages nécessaires afin d'optimiser au mieux le résultat final et de diminuer l'apparition de situations de conflit.

### **Théorie des ensembles flous**

La théorie des ensembles flous a été introduite par Zadeh en 1965 afin de tenir compte explicitement des informations imprécises au sein d'un système de décision. Dans le cadre de la théorie des ensembles stricts (ensembles classiques), un élément n'a que deux possibilités : soit il appartient à cet ensemble, soit il n'y appartient pas. C'est ce que l'on

appelle couramment de la logique binaire. On définit une fonction d'appartenance binaire  $\mu_A(x)$  entre un élément  $x$  et un ensemble  $A$  comme étant :

$$\mu_A(x) = 1 \text{ si } x \in A \quad (3.2)$$

$$= 0 \text{ si } x \notin A \quad (3.3)$$

En revanche, dans la théorie des ensembles flous, cette fonction d'appartenance ne prend pas seulement ces valeurs dans l'ensemble  $\{0,1\}$  mais dans tout l'intervalle  $[0,1]$ . Un sous-ensemble flou est donc défini de la manière suivante :

**Définition : 3.2** *Soit  $X$  un ensemble. Un sous-ensemble flou  $A$  de  $X$  est défini par une fonction d'appartenance  $f_A$  sur  $X$  à valeurs dans l'intervalle  $[0,1]$ .*

Ainsi un élément  $x$  appartiendra « un peu » à un ensemble  $A$  si la valeur de  $f_A(x)$  est proche de zéro, et a contrario, il appartiendra « beaucoup » à cet ensemble si sa fonction d'appartenance est forte (i.e. proche de 1). Un élément dont la fonction d'appartenance en ce point est nulle signifie qu'il n'appartient « pas du tout » à cet ensemble. Si en revanche la valeur est égale à 1, l'élément appartient complètement à cet ensemble.

Ces fonctions d'appartenance peuvent être fixées de manière totalement arbitraire ce qui pose le problème de la détermination de ces fonctions pour une application pratique réelle. Dans le cas de la fusion de données, on modélise donc l'imprécision et l'incertitude d'une donnée par un ensemble de fonctions d'appartenance à différentes classes, qui permettent de représenter les différentes hypothèses à définir. On définit ainsi des fonctions d'appartenance pour chaque élément provenant d'une source à fusionner, et on procède ensuite à des combinaisons.

Le raisonnement en logique floue s'appuie sur l'utilisation d'implication floue.

**Définition : 3.3** *Une implication floue est une relation  $R$  entre 2 ensembles  $X$  et  $Y$  quantifiant le degré de vérité de la proposition :*

$$SI (x \in A) \text{ ALORS } (y \in B)$$

avec  $A$  et  $B$  deux sous-ensembles flous respectivement de  $X$  et  $Y$ .

De manière similaire à la logique classique, on définit un ensemble de règles d'inférences reliant les entrées et les sorties du type :

$$SI (x \in A) \text{ ET } (y \in B) \text{ ALORS } (z \in C)$$

où  $A$ ,  $B$  et  $C$  sont des sous-ensembles flous.

Exemple : dans le cadre du maintien à domicile, on peut trouver les règles suivantes :

*SI position est allongée ET localisation est lit ALORS danger est faible*  
*SI position est allongée ET localisation est cuisine ALORS danger est fort*

Les fonctions d'appartenance des ensembles d'entrée et de sortie représentent l'intérêt principal de la théorie des ensembles flous. La détermination de ces fonctions d'appartenance repose souvent sur l'étude de données statistiques, des avis d'expert et des ajustements en fonction de campagne de tests antérieurs. La prise en compte d'entrées et de sorties multiples peut entraîner une explosion du nombre de règles qui alourdissent le fonctionnement global de l'approche.

**Remarque 3.10** *C'est une approche fréquemment utilisée pour des applications de maintien à domicile. Elle permet de rassembler des informations provenant de sources variées. Pour le dispositif EMUTEM, un algorithme de logique floue a été utilisé [Med09]. Les sorties des trois sous-systèmes ANASON, RFPat et Gardien constituent les entrées de l'algorithme qui a pour objectif de renseigner deux informations : une alerte en cas de situation de danger et la localisation de la personne. Dans notre application, la logique floue sera fort utile pour déterminer la posture d'une personne comme on le verra dans la partie consacrée aux résultats expérimentaux.*

### **Théorie des possibilités**

Introduite également par Zadeh en 1978, la théorie des possibilités repose aussi sur la théorie des sous-ensembles flous. Elle permet, comme la logique floue, de gérer l'imprécision, et contrairement à la logique floue, de gérer l'incertitude de certaines données. La formalisation des incertitudes sur des événements se fait de manière non probabiliste. L'objectif est de chercher à exprimer dans quelle mesure la réalisation d'un événement particulier est possible et à quel point on est certain de sa réalisation sans avoir connaissance de la probabilité que cet événement survienne. Deux mesures sont introduites :

- la possibilité qui traduit la consistance d'un événement par rapport à l'ensemble des croyances ;
- la nécessité qui traduit le degré avec lequel on attend l'occurrence d'un événement.

Il est alors possible d'associer à une source d'information  $X$  une distribution de possibilités noté  $\pi_X$  qui est une fonction définie sur  $X$  et qui prend ses valeurs dans  $[0,1]$  telle que pour tout  $x \in X$ ,  $\pi_X(x)$  traduit le degré de possibilité pour que  $X = x$ . Cette distribution de possibilité satisfait la condition de normalisation suivante :

$$\sup_{x \in X} \pi(x) = 1 \tag{3.4}$$



Le principe consiste donc à déterminer ces distributions de possibilités pour chaque source concernée et de les fusionner en choisissant parmi deux opérateurs fondamentaux en théorie des possibilités [Blo03] :

- opérateur conjonctif (intersection des distributions de possibilités) qui donne une idée pessimiste de la fusion ;
- opérateur disjonctif (union des distributions de possibilités) qui au contraire donne une vision optimiste de la combinaison.

Certains opérateurs s'adaptent en fonction de l'évolution de la situation observée. Selon les degrés de conflit et la fiabilité que l'on attribue aux différentes sources, l'opérateur opte pour un comportement conjonctif, disjonctif ou un compromis entre les deux. Le choix dépend également de ce que l'on recherche : entre une information certaine mais imprécise ou un résultat précis mais sûrement faux.

Comme pour la logique floue, une décision doit être prise suite à la combinaison des distributions de possibilités. Cette prise de décision se fait généralement à l'aide de deux critères : le critère du maximum et le critère de seuil qui fournissent respectivement une solution unique ou une plage de valeurs possibles (précision du résultat).

La différence majeure entre la théorie des possibilités et la théorie des probabilités est de faire la distinction entre incertitude et imprécision, ce qui est plus difficile avec des probabilités.

***Remarque 3.11*** *La théorie des possibilités constitue une perspective d'amélioration de notre dispositif de détermination de posture basé sur la logique floue. La prise en compte de l'incertitude permettrait probablement d'améliorer les performances globales de l'algorithme tout en s'appuyant sur la modélisation existante des fonctions d'appartenance.*

### **Autres approches dérivées de l'intelligence artificielle**

Plus récemment, des techniques dérivées de l'intelligence artificielle se sont développées. Parmi celles-ci, on retrouve les systèmes experts et les systèmes de fusion à haut niveau (systèmes à base de connaissance). Les règles d'inférences sont de la forme SI ...ALORS ...SINON. Chaque proposition se voit ainsi affectée d'un degré de confiance (nombre entre -1 et 1), obtenu en faisant inférer le système à base de règles portant sur des faits. Si une proposition a un degré de confiance égal à 1, alors on est sûr qu'elle est vérifiée. En revanche, si son degré de confiance est égal à -1, alors on est sûr qu'elle est fausse.

Le jugement est nuancé pour des valeurs intermédiaires. Le 0 est un cas particulier et exprime l'ignorance que l'on a sur une proposition.

**Remarque 3.12** *Ce type de système permet bien de définir l'incertitude mais n'est pas adapté pour la gestion de l'imprécision.*

**Remarque 3.13** *Des recherches sur la conception d'une chambre hospitalière intelligente ont été menées dans [Ria00]. A partir de valeurs physiques des capteurs (détecteurs IRP, microphone) et de variables sémiotique (durée, état du patient), un agent de perception identifie des faits et suivant la séquence de faits identifie un événement. Des règles portant sur les variables physiques et sémiotiques et les événements permettent la génération d'alarmes ou de messages.*

Une autre technique utilisée en fusion est la technique basée sur les réseaux de neurones. Élaborés dans les années 1970, les réseaux de neurones ont été développés à la base pour imiter le cerveau humain. Le réseau est constitué d'un ensemble de neurones qui possèdent des entrées et des sorties ainsi qu'une fonction d'activation. En fonction des entrées et de sa fonction d'activation, le neurone active une ou plusieurs sorties. On connecte ensuite un certain nombre de neurones entre eux et on présente à ce réseau un ensemble d'apprentissage. Chaque neurone adapte alors sa fonction d'activation en fonction de cet ensemble d'apprentissage. Une fois que le réseau a « appris », il est capable de classer un élément nouveau. Les réseaux de neurones sont intéressants dans le cas où l'expertise est délicate à formuler et où l'on possède un ensemble d'apprentissage conséquent. Ils se révèlent très utiles pour la modélisation des connaissances pour les systèmes basés sur la théorie des possibilités ou ceux basés sur la théorie de l'évidence.

**Remarque 3.14** *Il existe, à notre connaissance, très peu de systèmes d'habitat intelligent, basés entièrement sur les réseaux de neurones [Coo03]. Un inconvénient de ces méthodes est la relative lourdeur de mise en place et de maintien des bases de connaissances. De plus, la forte demande de puissance de calcul rend difficile pour l'instant, une application en temps réel.*

### 3.2.5 Conclusion sur la fusion de données

La fusion de données correspond donc à un procédé de combinaisons de données hétérogènes, le terme donnée devant être considéré dans son sens le plus large, c'est-à-dire à la fois les données provenant de capteurs mais également les informations issues de bases de connaissances, d'avis d'experts,... La conception d'un dispositif global de fusion de données nécessite que l'on se pose certaines questions avant tout développement. [Ruh07]

liste ces questions qui permettent de soulever les points importants à aborder en préambule de toute conception :

- Pourquoi et comment l'information est obtenue ?
- Sélection et définition de ou des quantités à mesurer ?
- Quels sont les capteurs qui me permettent de mesurer ces quantités ?
- Des connaissances a priori sont-elles nécessaires ?
- Existe-t-il des modèles fiables déjà disponibles ?
- Comment doivent être placés les capteurs ?
- Comment les signaux sont-ils transmis ?
- A quels moments doivent être effectuées les mesures ?

Au travers de ces questions, on parvient à mettre en avant les besoins matériels à mettre en place (réseaux étendus ou non, travail en ligne ou hors-ligne, traitement des signaux,...), à faciliter les choix des méthodes et des techniques selon le type de données que l'on va devoir manipuler et enfin à déterminer les modèles utiles. L'architecture du dispositif de fusion se dessine à partir de ces réflexions. Elle comporte généralement trois niveaux : la fusion de mesures, la fusion de primitives et la fusion de décisions.

La fusion de données n'est pas toujours la solution miracle pour résoudre un problème ce qu'illustre bien [Hal09] avec ce qu'il désigne comme les « Dirty Secrets in Data Fusion ».

1. Il n'y a pas de remplaçant à un bon capteur. Si quelque chose ne peut pas être observée ou déduite à partir des effets, alors aucune quantité de données fusionnées à partir de plusieurs capteurs ne pourra surmonter le problème.
2. Les traitements en aval ne peuvent corriger les méfaits de traitements en amont. A chaque étape, il est nécessaire d'utiliser le meilleur traitement possible.
3. La réponse issue de la fusion peut être pire que le meilleur capteur. Il faut correctement quantifier la précision de chaque capteur/entrée afin de ne pas créer de biais et d'erreurs dans le résultat de fusion.
4. L'algorithme magique n'existe toujours pas : aucun algorithme unique n'est la solution idéale en toutes circonstances.
5. Il n'y aura jamais assez de données d'apprentissage. Cependant les méthodes hybrides, combinant des informations explicites et implicites, peuvent aider.
6. Les recherches ont commencé du mauvais côté (c'est à dire du côté des capteurs et non du côté humain). Énormément de recherches ont été menées au niveau 0 et 1. Mais des recherches doivent être menées en commençant du côté humain (c'est à

dire la formulation d'hypothèses ou d'interprétation sémantique d'événements) d'où l'introduction du niveau 5.

7. Il est difficile de quantifier la valeur de l'apport de la fusion de données.
8. La fusion n'est pas un processus statique.

Bien conscients de ces difficultés, nous proposons, pour la problématique CAPTHOM, une architecture de fusion faisant intervenir l'ensemble des capteurs retenus et en découpant le traitement en quatre modules.

### 3.3 Architecture de fusion pour le projet CAPTHOM

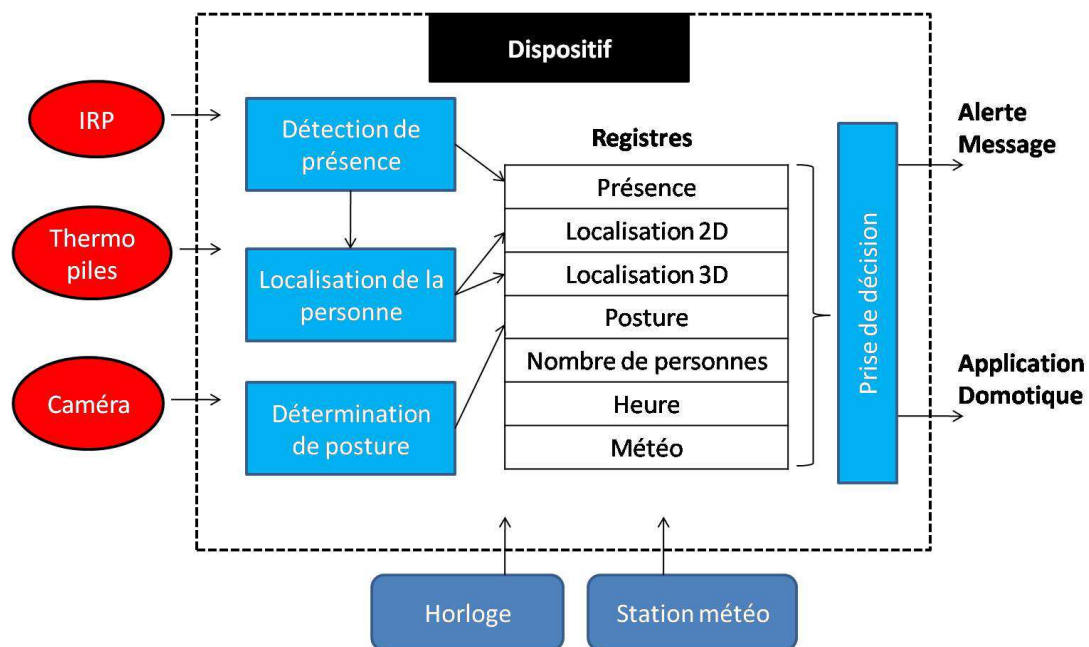


FIG. 3.9 – Schéma de l'architecture de fusion de données.

D'un point de vue général, le dispositif CAPTHOM peut être schématisé comme suit (figure 3.9) :

- En entrée : les capteurs choisis, à savoir des détecteurs IRP, des thermopiles et une caméra, auxquels s'ajoutent une horloge et une station météo ;
- En sortie : les prises de décision telles que les alertes et le déclenchement d'applications domotiques respectivement pour le maintien à domicile et la gestion énergétique du bâtiment.

L'architecture de fusion de données représente le cœur du dispositif. Elle est constituée de quatre nœuds ou modules de fusion. Ces modules de fusion sont : (1) la détection de présence qui utilise les données provenant des trois types de capteurs, (2) la localisation qui utilise les données visuelles à travers le VRHE et les détecteurs IRP à travers un zonage de l'environnement, (3) la détermination de posture qui utilise des informations extraites de l'image, (4) un outil de prise de décision pour l'application visée.

Dans la suite, l'organigramme d'un module de fusion utilise le formalisme suivant : les capteurs sont symbolisés par des ovales rouges, les méthodes par un rectangle blanc, les informations en sortie par des parallélogrammes bleus.

### **3.3.1 Fusion pour la détection de présence humaine**

Pour le module de détection de présence, le rôle de chacun de capteurs choisis est le suivant :

- le détecteur IRP détecte un mouvement ;
- la thermopile détecte une présence ;
- la caméra détecte la présence et l'identifie.

Comme on l'a déjà mentionné, ces technologies sont sensibles aux bruits et à des perturbations. Nous utilisons la complémentarité et la redondance de ces informations en fusionnant les données pour assurer une détermination la plus fiable possible de la présence humaine à tout instant.

#### **3.3.1.1 Processus de détection de présence humaine**

Le fonctionnement du module de fusion pour la détection humaine est représenté par la figure 3.10. Pour des raisons énergétiques évidentes, la caméra ne fonctionne pas tout le temps. Elle est mise en marche si le détecteur IRP ou la thermopile détecte respectivement un mouvement ou une présence. La caméra confirme ou infirme la présence (robustesse face aux fausses détections). Ensuite, l'entité détectée est classifiée grâce à l'algorithme de traitement d'images en tant qu'entité humaine ou non. L'algorithme de traitement d'images attribue à chaque personne détectée un numéro d'identification. Ainsi, on récupère également l'information concernant le nombre de personnes présentes dans la pièce. La détection de présence humaine est une fusion de décisions.

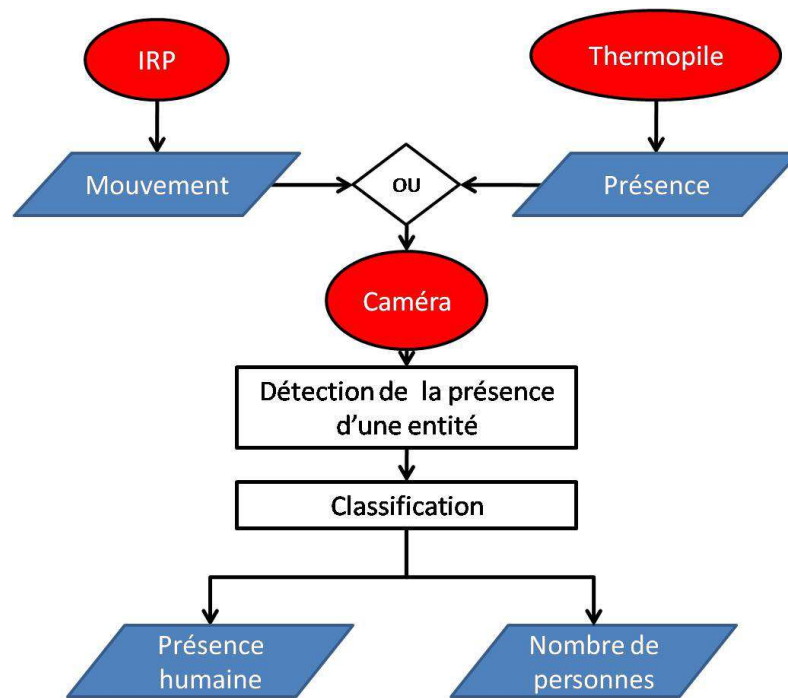


FIG. 3.10 – Organigramme de la détection de présence humaine.

**Remarque 3.15** Si la caméra est interdite dans certaines pièces (toilettes, salle de bains) pour respecter la vie privée de la personne, on privilégie la combinaison des deux capteurs détectant le rayonnement thermique (détecteur IRP et thermopile). La thermopile, dans des zones de petites dimensions, assure une réelle détection de présence, mais ne la caractérise pas.

### 3.3.1.2 Exemples d'application

#### Simulation dans une pièce

Afin d'illustrer le processus décrit précédemment, nous avons défini, à l'aide du logiciel SNOOPS, un environnement dans lequel sont positionnés trois capteurs : un détecteur IRP, une thermopile et une caméra. Les trois capteurs sont placés aux mêmes coordonnées et orientés suivant le même axe. Ce choix a été fait afin de mettre en avant les propriétés de redondance et de complémentarité de ces trois technologies pour une application de détection de présence.

**Remarque 3.16** Les caractéristiques des capteurs (angle d'ouverture, portée) sont celles que l'on retrouve dans les dispositifs commercialisés.

La thermopile possédant une portée relativement faible (sensibilité jusqu'à 2m), les

dimensions de l'environnement ont été choisies de telle sorte que la réponse de cette dernière soit exploitable. L'environnement et les zones de détection de chaque capteur sont présentés à la figure 3.11. La zone délimitée en rouge est associée au détecteur IRP, la zone bleue correspond à la zone de détection de la thermopile et la zone verte représente la zone couverte par la caméra. Un obstacle est présent au centre de la pièce et obstrue le champ de vision des trois détecteurs. Le point noir correspond à la position de la personne.

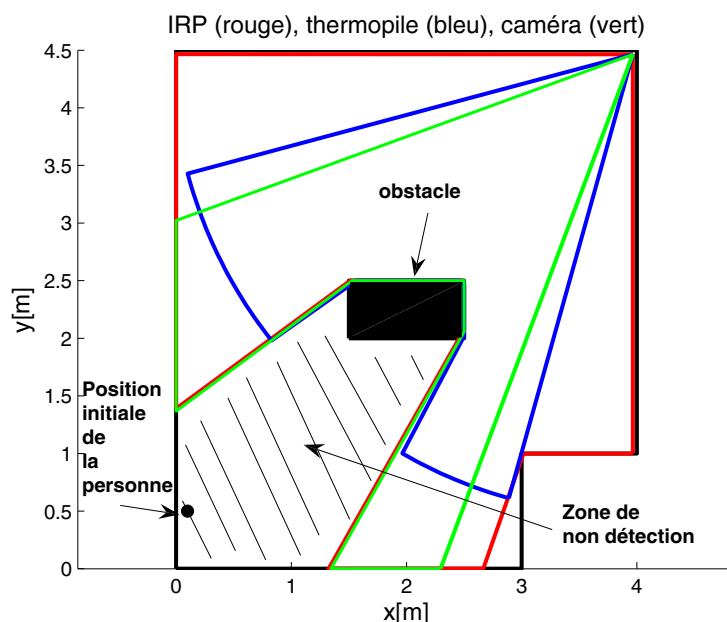


FIG. 3.11 – Placement et couverture des capteurs pour le scénario N°1.

**Remarque 3.17** *Le placement des capteurs n'est pas optimisé dans cette configuration (présence d'une large zone non couverte par les capteurs).*

On fait évoluer la position de la personne au cours du temps en ajoutant une information supplémentaire : son activité (1 si elle est immobile, 2 si elle est en mouvement).

Le premier scénario considéré est décrit par le tableau 3.2 et illustré à la figure 3.12.

**Remarque 3.18** *Les positions définies dans le tableau 3.2 constituent les principales étapes du scénario. Afin de travailler avec plus de points, nous effectuons une discrétisation entre deux étapes afin d'obtenir des positions intermédiaires.*





seuil signifie qu'une source chaude est présente dans le cône de détection du capteur.

- La caméra offre la possibilité de caractériser l'entité détectée par l'IRP et/ou la thermopile comme étant humaine ou non. Deux sorties sont considérées : 1 si l'entité est humaine, 0 si l'entité n'est pas humaine ou si le capteur ne détecte pas de présence.

En sortie, la décision concernant la présence se décline selon quatre valeurs : (0) absence, (1) mouvement, (2) présence, (3) présence humaine.

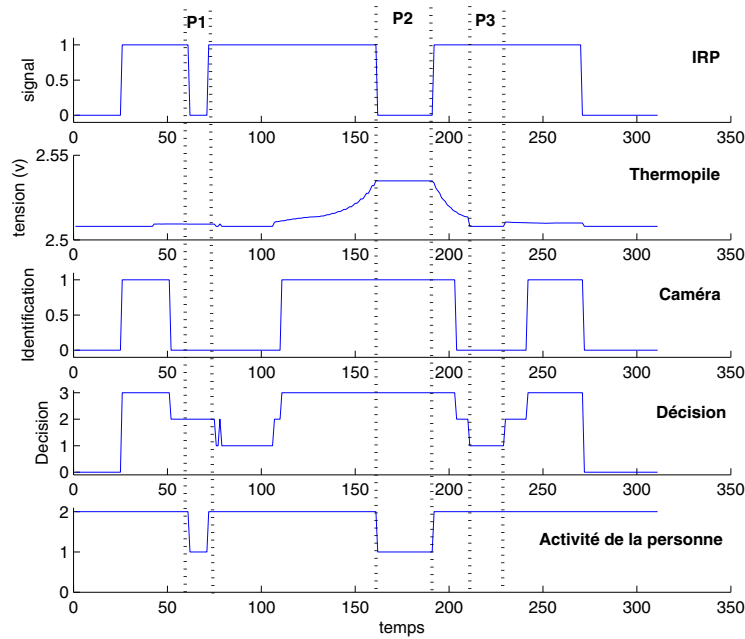


FIG. 3.13 – Réponse des capteurs et décision finale pour le scénario N°1.

### La prise de décision

Afin de statuer sur la présence ou non d'une personne dans l'environnement surveillé, nous utilisons une méthode basée sur la théorie des ensembles. Des règles d'inférence faisant intervenir les réponses des capteurs à chaque instant sont donc définies pour aider à la prise de décision.

Dans l'espace de temps  $P1$ , la personne est immobile. Le détecteur IRP ne voit donc pas la personne. En revanche, la thermopile détecte une présence dans son champ de vision, le signal est toutefois très faible étant donné la distance à laquelle se trouve la personne par rapport au capteur. La thermopile envoie donc un message à la caméra qui

délivre en sortie la valeur 0. Ceci s'explique par le fait que la personne se trouve hors du champ de vision de la caméra. La décision finale est donc la présence d'une entité, humaine ou non.

La personne est de nouveau immobile pendant la période  $P2$ . Elle s'est toutefois rapprochée des capteurs comme le montre l'évolution de la tension délivrée par la thermopile. Comme pendant  $P1$ , l'IRP ne détecte rien et la thermopile détecte une présence. Elle interroge donc la caméra qui cette fois-ci voit une entité dans l'image et la caractérise comme étant humaine.

La personne s'éloigne ensuite des capteurs et sort des champs de vision respectifs de la thermopile et de la caméra. Pendant la période  $P3$ , seul le détecteur IRP détecte le mouvement de la personne. La décision finale est donc une détection de mouvement.

En résumé, pendant les phases de mouvement, le détecteur IRP interroge la caméra afin qu'elle confirme que l'entité en mouvement est bien une personne. Il peut arriver en effet que l'IRP détecte le mouvement d'un animal ou soit perturbé par un flux thermique provoqué par un appareil de chauffage ou un courant d'air. La thermopile délivre une information redondante qui confirme la présence d'une entité. Nous avons choisi l'IRP pour « réveiller » la caméra dans la mesure où il possède un champ de vision plus vaste que la thermopile et permet ainsi de détecter la personne dès l'entrée dans la pièce. En revanche, pendant les phases d'immobilité, c'est la thermopile qui active la caméra. Une difficulté peut survenir si la période d'immobilité se prolonge.

Afin d'illustrer ce qui se passe lors de longues périodes d'immobilité, par exemple une personne regardant la télévision, nous définissons un nouveau scénario dans lequel nous introduisons une phase d'immobilité plus importante. Les réponses des capteurs et la décision finale pour le scénario N°2 sont présentées à la figure 3.14.

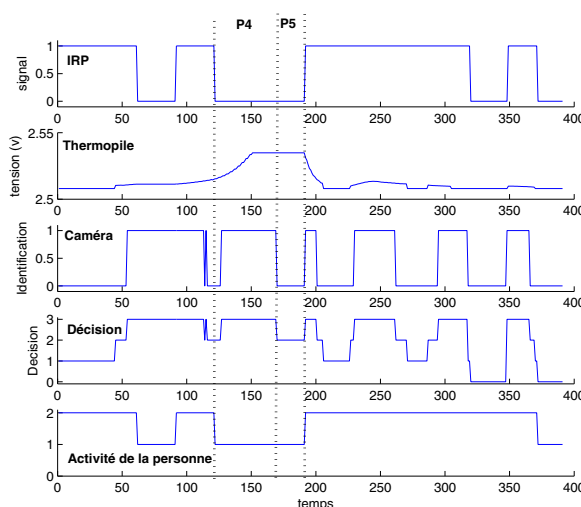


FIG. 3.14 – Réponse des capteurs et décision finale pour le scénario N°2.

La période  $P4$  est similaire à la période  $P2$  du scénario N°1. Le dispositif conclut donc qu'une personne est présente dans la pièce. En revanche, à l'inverse de  $P2$ , la personne reste immobile plus longtemps. Par conséquent, elle finit par être confondue avec l'arrière plan et n'est donc plus détectée lors de la phase de soustraction de l'arrière plan du traitement d'images. La caméra se remet donc en état de veille. Tant que la personne reste immobile (période  $P5$ ), seule la thermopile détecte une présence. La décision finale bascule donc d'une détection de présence humaine à une détection de présence. Il est toutefois envisageable de maintenir la sortie sur présence humaine dans la mesure où une personne a été détectée lors du précédent pas d'échantillonnage.

En prenant en compte cet effet mémoire, on obtient donc pour le scénario N°2 la prise de décision illustrée à la figure 3.15.

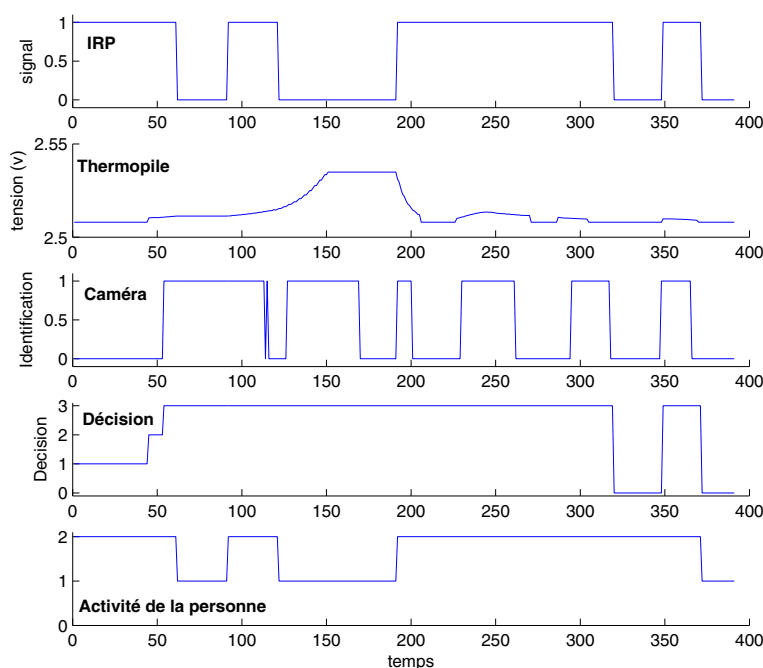


FIG. 3.15 – Réponse des capteurs et décision finale pour le scénario N°2 avec mémorisation.

La figure 3.16 permet d'illustrer le comportement du dispositif de détection de présence en cas de mouvement d'un animal dans le champ de détection des capteurs (fausses détections). La période  $P6$  correspond à l'entrée d'un animal de grande taille (un chien) dans la pièce. L'animal est détecté par le détecteur IRP ainsi que par la thermopile. Cependant l'algorithme de traitement d'images caractérise la cible en mouvement comme

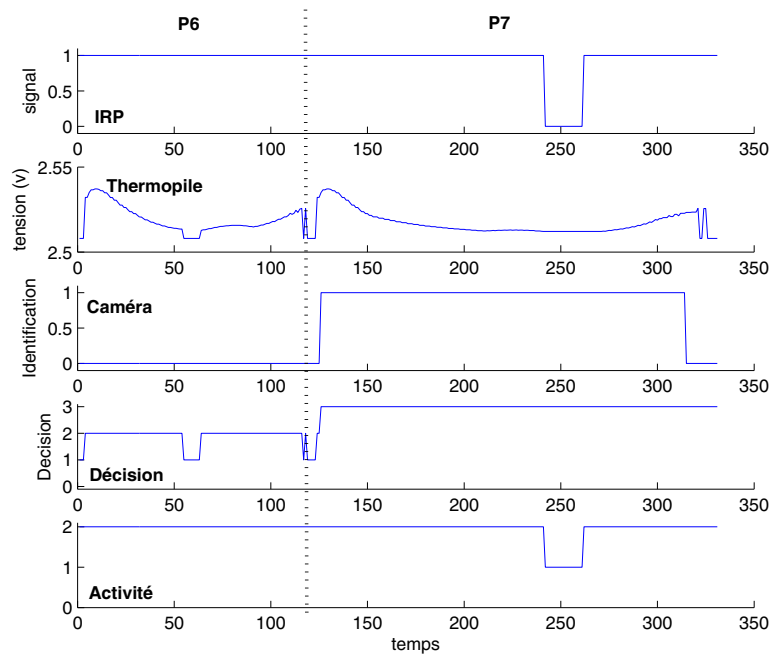


FIG. 3.16 – Réponse des capteurs et décision finale pour le scénario N°3 avec mémorisation.

étant non humaine. On maintient donc une décision de mouvement ou de présence selon que l'animal est dans le champ de détection de la thermopile. En  $P7$ , une personne entre dans la pièce. La décision de présence est alors modifiée en présence humaine.

**Remarque 3.19** *En condition réelle, la réponse de la thermopile lors de la présence d'un animal de grande taille est différente de la réponse traduisant la présence d'un humain à la même distance du capteur.*

### 3.3.2 La localisation 2D et 3D

Les exigences de la localisation 2D et 3D varient en fonction de l'application. La localisation 3D peut être peu précise pour une application domotique contrairement à une application de maintien à domicile.

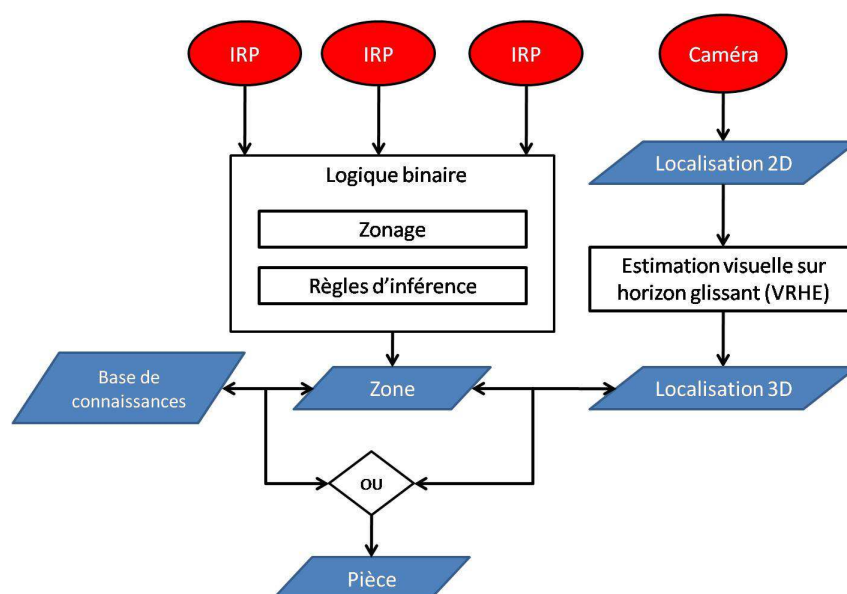


FIG. 3.17 – Organigramme de détermination de la localisation.

Le module de localisation 2D/3D est représenté par la figure 3.17. Dans la mesure où il n'est pas toujours possible d'équiper l'habitat à l'aide de caméra, une localisation 3D reste envisageable en utilisant plusieurs détecteurs IRP disposés dans l'habitat de manière à maximiser les zones de recouvrement des capteurs. On obtient ainsi un découpage plus ou moins fin de l'habitat. En se basant sur la théorie des ensembles et avec la définition de règles d'inférence portant sur les réponses des capteurs, on est capable de dire dans quelle zone, plus ou moins réduite, se trouve la personne.

Or, comme nous le verrons à travers un exemple, différentes pièces de l'habitat peuvent contenir des zones ayant le même identifiant. On ne peut donc pas en déduire directement la pièce dans laquelle se trouve la personne. Deux approches peuvent être envisagées : soit la prise en compte de la localisation 3D délivrée par une caméra, soit en gardant en mémoire la zone (ou la pièce) où se trouvait la personne à l'instant d'échantillonnage précédent.

La localisation 2D est directement obtenue par l'algorithme de détection de présence par traitement d'images.

### 3.3.2.1 La localisation 3D par zone

L'appartement témoin du GIS MADONAH a été découpé en zones, surveillées par un ou plusieurs détecteurs IRP. Leur placement a été optimisé à l'aide du logiciel de simulation SNOOPS pour maximiser les zones de recouvrement. En utilisant cinq détecteurs IRP, on obtient une couverture quasi-totale de l'habitat et un découpage suivant 15 zones, résumé par le tableau 3.3 et illustré par la figure 3.18, chaque zone faisant intervenir de un à quatre détecteurs.

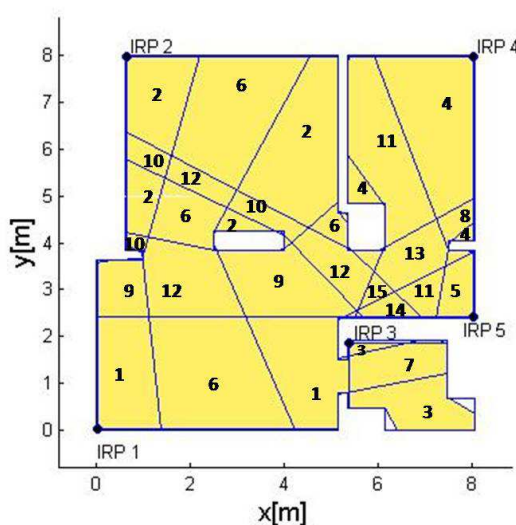


FIG. 3.18 – Découpage en zones de l'appartement témoin du GIS MADONAH.

zone	IRP1	IRP2	IRP3	IRP4	IRP5	zone	IRP1	IRP2	IRP3	IRP4	IRP5
1	×					9	×				×
2		×				10		×			×
3			×			11				×	×
4				×		12	×	×			×
5					×	13	×			×	×
6	×	×				14		×		×	×
7	×		×			15	×	×		×	×
8	×			×							

TAB. 3.3 – Liste des zones définies par les détecteurs IRP.

**Remarque 3.20** Les détecteurs IRP possèdent tous les mêmes caractéristiques. Ils sont positionnés sur le mur à hauteur de plafond et caractérisés par un angle d'ouverture de

90 degrés et un rayon de détection de 10m. Ces valeurs correspondent aux caractéristiques classiques des détecteurs IRP commercialisés.

Nous avons ensuite simulé différents scénarios comprenant des phases de déplacement et d'immobilité (figure 3.19). Les cercles bleus représentent l'emplacement des détecteurs IRP et les croix vertes les étapes successives du scénario. Le tableau 3.4 décrit un exemple de scénarios en précisant la position et l'activité de la personne à chaque phase du scénario (la personne se trouve initialement dans la zone 9, l'entrée de l'appartement (étape 1)).

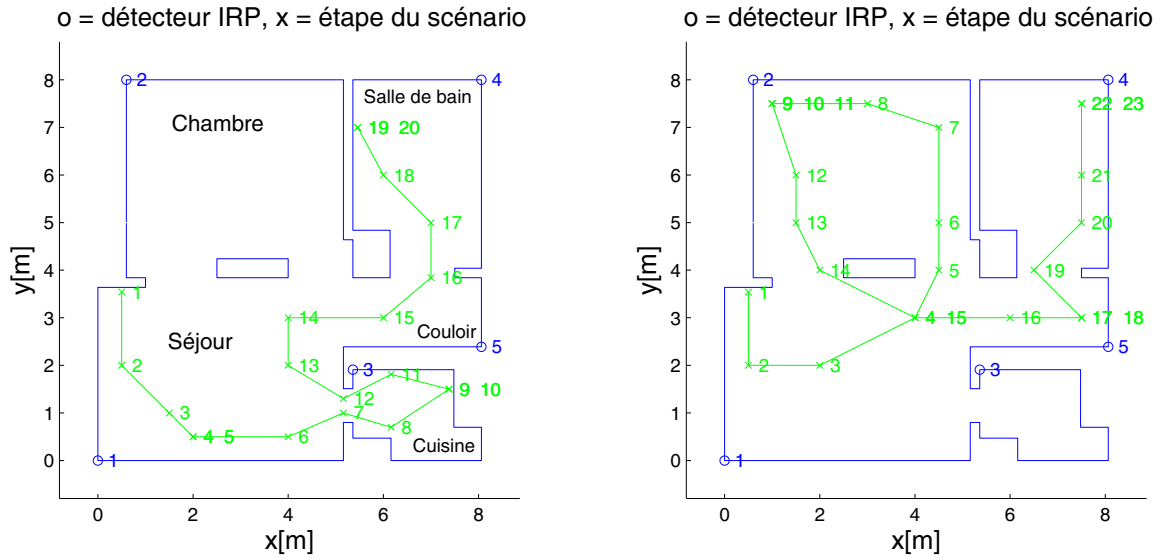


FIG. 3.19 – Placement des détecteurs IRP et exemples de scénarios simulés.

L'outil SNOOPS délivre alors les réponses de chacun des détecteurs à chaque étape du scénario. Une fusion de mesures permet de localiser la personne. Cette localisation est effectuée à l'aide d'une méthode basée sur la théorie des ensembles. Des règles d'inférence portant sur les états des capteurs sont définies. Un exemple de règle d'inférence est :

SI (IRP1 est 1) ET (IRP2 est 1) ET (IRP5 est 1) ALORS (localisation est Z12).

Pour le scénario décrit précédemment, la fusion des capteurs détermine la zone dans laquelle se trouve la personne (figure 3.20).

Plusieurs zones possèdent le même numéro d'identification ce qui implique que l'on ait besoin d'informations supplémentaires pour diminuer l'incertitude. Si une personne est détectée en zone 6, elle peut être soit dans le séjour soit dans la chambre. Pour lever l'ambiguïté, deux approches peuvent être envisagées. Par exemple, si l'on dispose de caméra, on compare l'estimation de position obtenue à l'aide du VRHE avec les ensembles

Étape	Position	Activité	Étape	Position	Activité
1	[0.5 3.54]	Normal	11	[6.16 1.81]	Normal
2	[0.5 2]	Normal	12	[5.16 1.3]	Normal
3	[1.5 1]	Normal	13	[4 2]	Normal
4	[2 0.5]	Normal	14	[4 3]	Normal
5	[2 0.5]	Immobile	15	[6 3]	Normal
6	[4 0.5]	Normal	16	[7 3.84]	Normal
7	[5.16 1]	Normal	17	[7 5]	Normal
8	[6.16 0.7]	Normal	18	[6 6]	Normal
9	[7.38 1.5]	Normal	19	[5.46 7]	Normal
10	[7.38 1.5]	Immobile	20	[5.46 7]	Immobile

TAB. 3.4 – Exemple de scénario simulé à l’aide de l’outil SNOOPS.

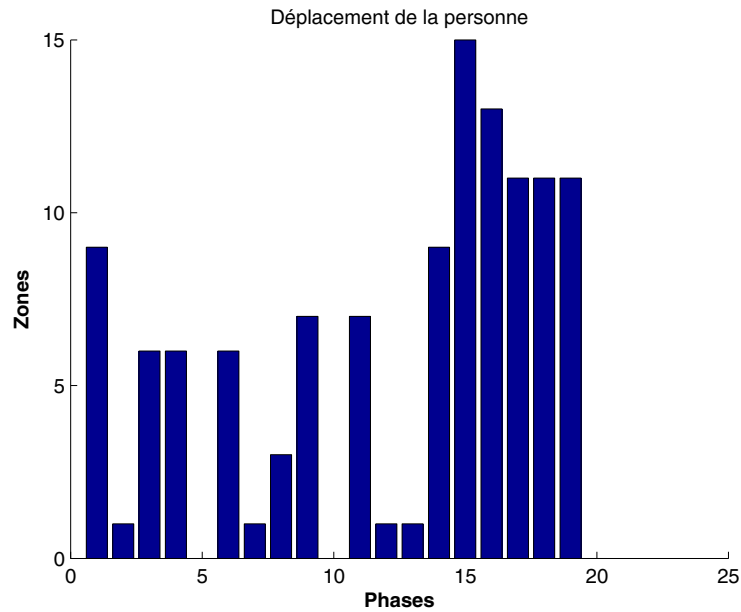


FIG. 3.20 – Résultat obtenu pour le scénario du tableau 3.4.

de coordonnées des zones considérées. A l’inverse, si l’on ne dispose pas de capteurs supplémentaires, on utilise l’information passée. On regarde dans quelle zone ou dans quelle pièce se trouvait la personne au pas d’échantillonnage antérieur.

Si l’on effectue cette dernière démarche sur le scénario du tableau 3.4, on obtient les localisations suivantes :



Étape	Zone	Décision		Étape	Zone	Décision
1	9	séjour		11	7	cuisine
2	1	séjour		12	1	séjour
3	6	séjour		13	1	séjour
4	6	séjour		14	9	séjour
5	0	séjour		15	15	couloir
6	6	séjour		16	13	couloir
7	1	séjour		17	11	couloir ou salle de bain
8	3	cuisine		18	11	couloir ou salle de bain
9	7	cuisine		19	11	couloir ou salle de bain
10	0	cuisine		20	11	couloir ou salle de bain

La zone 0 signifie qu'aucun capteur ne détecte un mouvement ce qui arrive notamment lorsque la personne reste immobile. Dans ce cas, on considère que la personne est encore dans la zone où elle se trouvait au pas précédent. La mémorisation de l'état précédent permet donc de gérer à la fois les situations d'ambiguïté et les périodes d'immobilité.

Il subsiste une incertitude pour les quatre dernières étapes du scénario. Ceci est dû à l'impossibilité de modéliser la fermeture de la porte qui sépare le couloir de la salle de bain. En condition réelle, seul le capteur IRP4 est en mesure de détecter la personne dans la salle de bain.

***Remarque 3.21*** *Pour des petites pièces ou des zones restreintes, l'utilisation d'une thermopile apporte une information redondante qui permet de lever les dernières incertitudes.*

On considère un second scénario (vignette droite de la figure 3.19) comportant quatre phases d'immobilité (étape 10, 11, 18 et 23). Les résultats de localisation issues de la fusion sont donnés à la figure 3.21.

En utilisant la démarche décrite précédemment, on détermine les localisations suivantes :

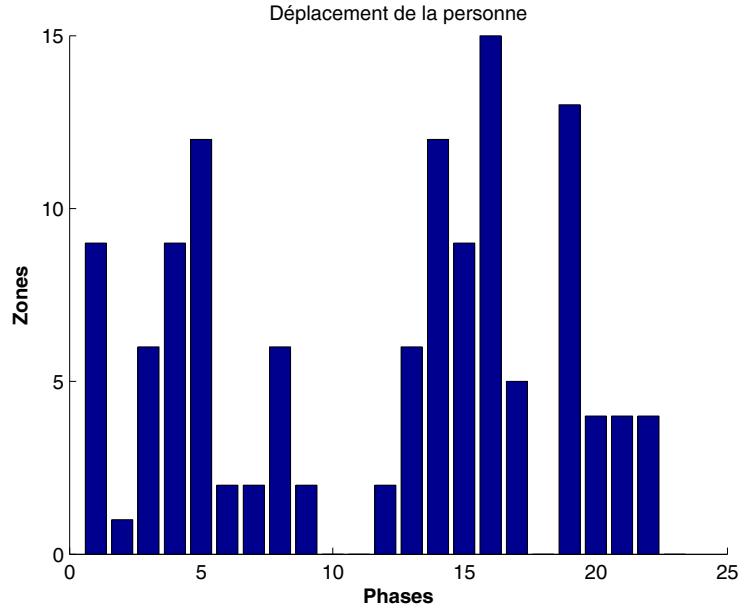


FIG. 3.21 – Résultat obtenu pour un second scénario.

Étape	Zone	Décision		Étape	Zone	Décision
1	9	séjour		13	6	chambre
2	1	séjour		14	12	séjour ou chambre
3	6	séjour		15	9	séjour
4	9	séjour		16	15	couloir
5	12	séjour ou couloir		17	5	couloir
6	2	chambre		18	0	couloir
7	2	chambre		19	13	couloir
8	6	chambre		20	4	salle de bain
9	2	chambre		21	4	salle de bain
10	0	chambre		22	4	salle de bain
11	0	chambre		23	0	salle de bain
12	2	chambre				

### 3.3.2.2 La localisation 3D précise

Pour affiner la localisation 3D, on utilise des informations plus riches, à savoir les images. L'estimation et le suivi de position de la personne dans l'image est réalisée à partir des sorties de l'algorithme de traitement d'images et à l'aide du VRHE.

**Remarque 3.22** *Les estimations issues de l'algorithme VRHE peuvent parfois être erro-*

nées à cause d'un trop grand changement de luminosité ou de phénomènes d'ombres. Les détecteurs IRP et les thermopiles permettent alors de corriger ces erreurs d'estimations dans la mesure où ils ne sont pas perturbés par ces difficultés. Les 3 types d'informations sont alors fusionnés.

Nous avons reproduit, sous SNOOPS, le placement des détecteurs IRP présents au sein de l'appartement témoin du GIS MADONAH (8 capteurs sont présents au total, nous n'en conservons que 2 pour notre exemple). Nous positionnons également une caméra ainsi qu'une thermopile (figure 3.22).

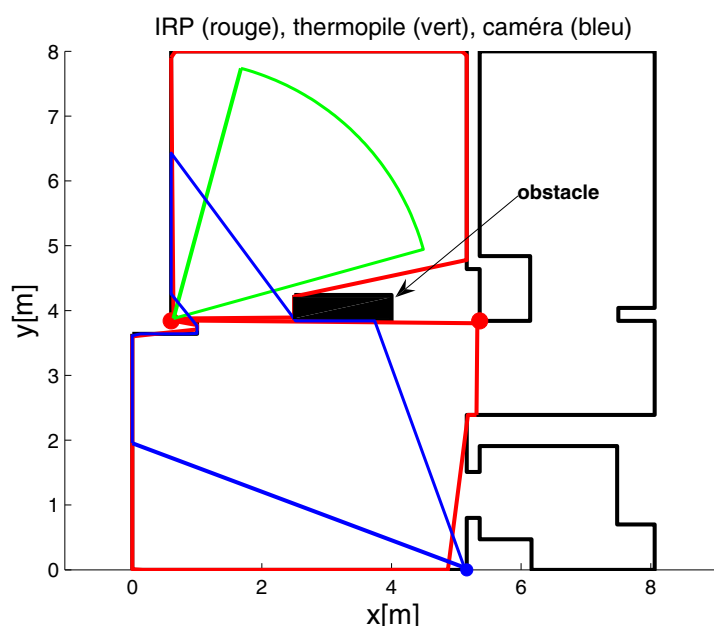


FIG. 3.22 – Placement et couverture des capteurs pour la localisation.

On simule un déplacement dans le salon de l'appartement (figure 3.23) ainsi que l'estimation de position 3D effectuée par la caméra à l'aide de l'approche VRHE. On suppose qu'au cours du déplacement, la caméra rencontre des difficultés pour estimer la position (occlusion, variations lumineuses), représentées par les points aberrants  $P_a b$ . Les détecteurs IRP et la thermopile corrigent alors cette estimation.

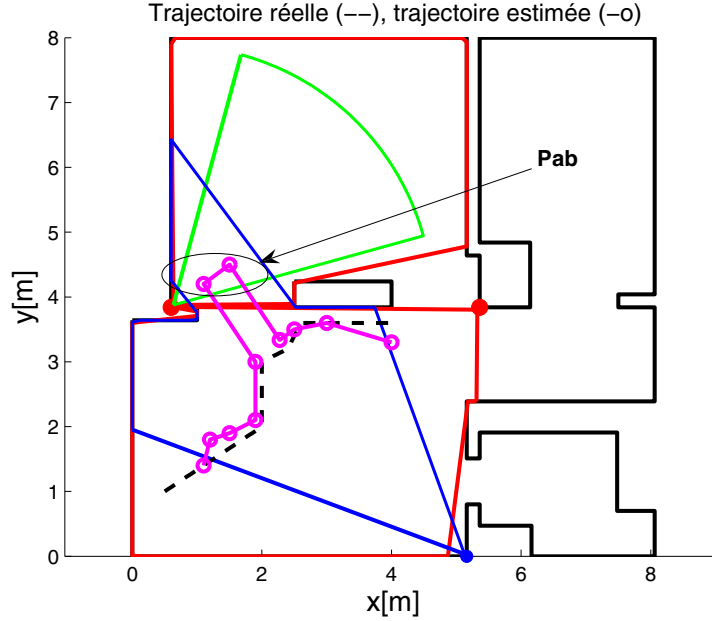


FIG. 3.23 – Trajectoire réelle (---), trajectoire estimée (-o).

### 3.3.3 Fusion pour la détermination de posture

Pour le maintien à domicile de personnes âgées, la détermination de la posture est une information très importante et très utile pour la prise de décision. Elle permet, entre autres, de détecter des situations d'alerte comme des chutes en associant la posture de la personne avec sa localisation dans l'environnement surveillé.

Beaucoup d'applications de surveillance de personnes âgées utilisent des capteurs de type accéléromètre ou gyromètre, embarqués sur la personne [Dem02][Med09]. Au regard des attentes et des spécifications des partenaires industriels du projet CAPTHOM, l'utilisation d'un dispositif porté par la personne n'est pas envisagée, la personne devant être complètement passive vis à vis du système (dispositif non invasif).

Nous avons donc opté pour une exploitation des données provenant du traitement d'images pour réaliser la détermination de la posture de la personne détectée. Nous rappelons qu'aucune image n'est visualisée à l'extérieur.

L'idée est de mettre en place une méthode qui offre la possibilité de fusionner différentes caractéristiques issues du traitement d'image pour déterminer la posture de la personne. Dans [Gir05], Girondel et al. décrivent un dispositif automatisé qui reconnaît automatiquement quatre postures statiques de l'homme pour la surveillance de personnes

âgées. Leur approche, basée sur l'utilisation de la théorie des croyances, offre de bonnes performances en terme de reconnaissance mais comporte certaines limitations. La détermination de la posture se fait par comparaison de trois caractéristiques à l'instant courant par rapport à une posture de référence. La personne doit toujours se tenir à la même distance de la caméra. De plus, un inconvénient de la théorie des croyances est la difficulté que l'on peut rencontrer lors de la détermination des modèles de croyance qui nécessite de nombreuses données expérimentales et/ou des avis d'expert.

Pour s'affranchir de cette dépendance vis à vis d'une posture de référence, nous proposons d'utiliser quatre paramètres obtenus par traitement d'images sur l'image courante. Vu le caractère imprécis et les possibles situations de conflit qui peuvent apparaître lors de la détermination d'une posture par rapport à une autre, le choix de la méthode de fusion s'est orienté vers une méthode basée sur une approche symbolique plutôt que probabiliste où l'imprécision des données est difficilement quantifiable. Les quatre paramètres sont fusionnés par une méthode basée sur la logique floue (figure 3.24). Les fonctions d'appartenance ont été définies grâce à des expérimentations.

Le processus de fusion pour la détermination de la posture peut être décrit par :

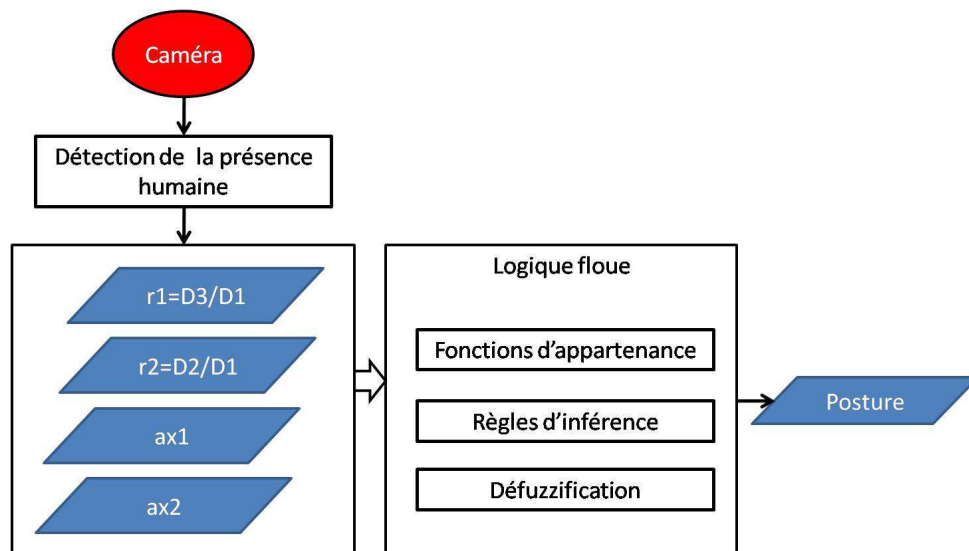


FIG. 3.24 – Organigramme de la détermination de la posture.

### 3.3.3.1 Méthodologie

Pour rappel, nous utilisons une seule caméra perspective, positionnée à une hauteur d'environ 2 mètres et permettant une couverture quasi-globale de la pièce. Nous faisons l'hypothèse que la personne est seule dans la pièce et qu'aucun obstacle n'obstrue le champ de vision de la caméra.

#### Les critères d'entrée

Les informations sont extraites grâce au traitement d'images. Dès l'étape de soustraction d'arrière plan, on dispose d'un ensemble de pixels blancs correspondant aux changements entre l'image courante et l'arrière plan, à savoir la personne détectée dans l'image (figure 3.25). Deux données sont alors extraites de ce nuage de points :

- $C$ , correspondant au point moyen (assimilable au centre de gravité, noté  $cdg$ ) ;
- $e$ , un des axes principaux de l'ensemble des points.

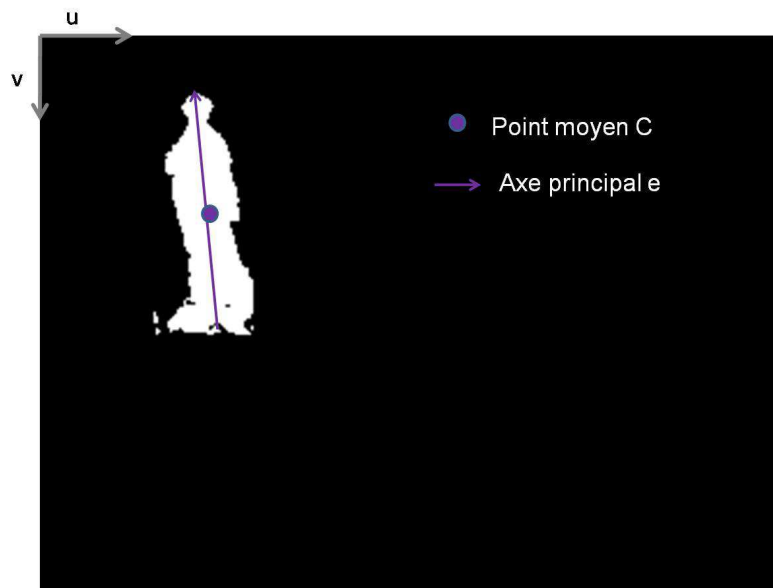


FIG. 3.25 – Représentation du point moyen et de l'axe principal du nuage de points.

**Remarque 3.23** *La détermination de  $C$  et de l'axe principal  $e$  est décrite dans l'annexe B.*

A partir de la boîte englobante, résultat final de l'algorithme de traitement d'images (figure 3.26), on mesure trois distances :

- $D1$ , la hauteur de la boîte englobante (assimilable en position debout à la distance tête/pieds) ;

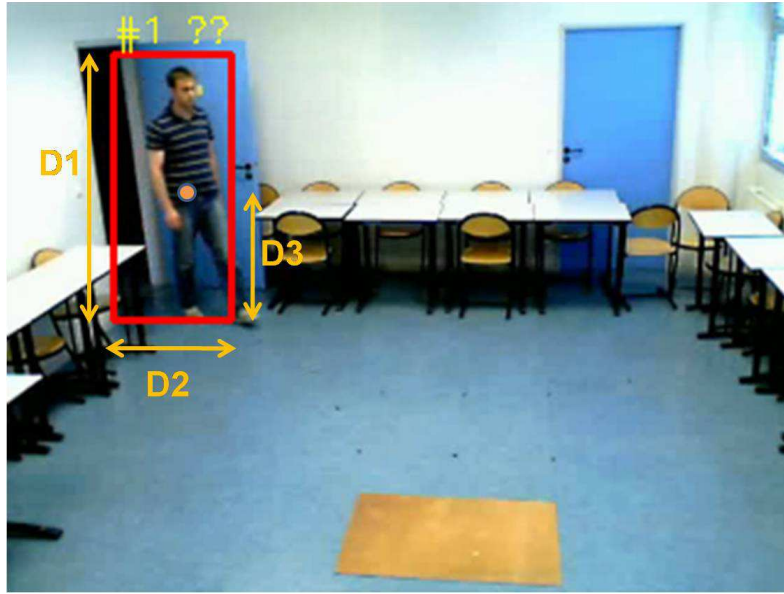


FIG. 3.26 – Représentation des distances  $D1$ ,  $D2$  et  $D3$ .

- $D2$ , la largeur de la boîte englobante ;
- $D3$ , la distance entre le point  $C$  (centre de gravité) et le bas de la boîte englobante.

A partir de ces données ( $D1$ ,  $D2$ ,  $D3$ ,  $C$  et  $e$ ), on définit quatre paramètres, utilisés en entrée du dispositif de logique floue :

- $r_1 = D3/D1$ , rapport entre la distance cdg/pied et la hauteur de la boîte ;
- $r_2 = D2/D1$ , rapport entre la largeur et la hauteur de la boîte ;
- $ax_1$ , la composante suivant  $u$  de l'axe principal  $e$  ;
- $ax_2$ , la composante suivant  $v$  de l'axe principal  $e$ .

**Remarque 3.24** *Il est important de noter que l'intérêt d'utiliser des rapports de distances réside dans la volonté de s'affranchir au maximum de la distance à laquelle se trouve la personne par rapport au dispositif visuel et de permettre la détermination de la posture quelle que soit la physiologie de la personne.*

Le dispositif de reconnaissance de posture, que nous proposons, est basé sur la fusion des paramètres  $r_1$ ,  $r_2$ ,  $ax_1$  et  $ax_2$  en utilisant un algorithme de logique floue. On opère une fusion de primitives.

## L'algorithme de logique floue [Kli95]

Nous avons identifié ci-dessus les ensembles d'entrées qui nous permettent d'identifier la posture de la personne. La mise en place d'un algorithme de logique floue nécessite la définition de l'univers de sortie qui, pour notre application, correspond à la posture de la personne. Cet univers de sortie est composé des sous-ensembles flous associés aux différentes postures considérées : couché ( $H_1$ ), accroupi ( $H_2$ ), assis ( $H_3$ ) et debout ( $H_4$ ). Étant donné que les quatre hypothèses de posture ne couvrent pas l'intégralité de l'univers de sortie, nous ajoutons une cinquième hypothèse  $H_0$  pour les postures indéterminées.

### (a) Les fonctions d'appartenance

Pour chaque entrée et pour la sortie, nous identifions les sous-ensembles flous pris en compte, chaque sous-ensemble étant défini par une fonction d'appartenance. La fonction d'appartenance d'un sous-ensemble flou caractérise le degré d'appartenance d'une variable dans le sous-ensemble considéré. A la différence de la théorie classique des ensembles, où un élément appartient à un unique ensemble, en logique floue, un même élément peut appartenir à plusieurs ensembles avec pour chacun un degré d'appartenance. Une fonction d'appartenance d'un sous-ensemble flou  $A$  d'un univers  $U$  est définie de la manière suivante :

$$\forall x \in U, f_A(x) \in [0; 1] \quad (3.5)$$

Les fonctions d'appartenance ont été définies à partir d'expérimentations. Des scénarios, illustrant la vie courante où chaque posture est considérée, ont été réalisés. Une analyse des séquences d'enregistrement a permis d'obtenir les courbes de variations des critères ( $r_1$ ,  $r_2$ ,  $ax_1$  et  $ax_2$ ). Un exemple est donné pour  $r_1$  et  $r_2$  avec deux scénarios (figures 3.27 et 3.28). En fonction de la sensibilité des critères par rapport à la posture, deux, trois ou quatre fonctions d'appartenance ont été retenues pour chaque critère. Pour  $r_2$  et  $ax_1$ , trois fonctions d'appartenance (valeur faible, moyenne et forte) sont utilisées (figure 3.29).

Pour  $r_1$  (paramètre plus sensible) et pour  $ax_2$ , on définit respectivement quatre et deux fonctions d'appartenance.

### (b) Les règles de combinaison

Comme en logique binaire, il est possible d'appliquer des raisonnements sur les sous-ensembles qui peuvent prendre la forme suivante :



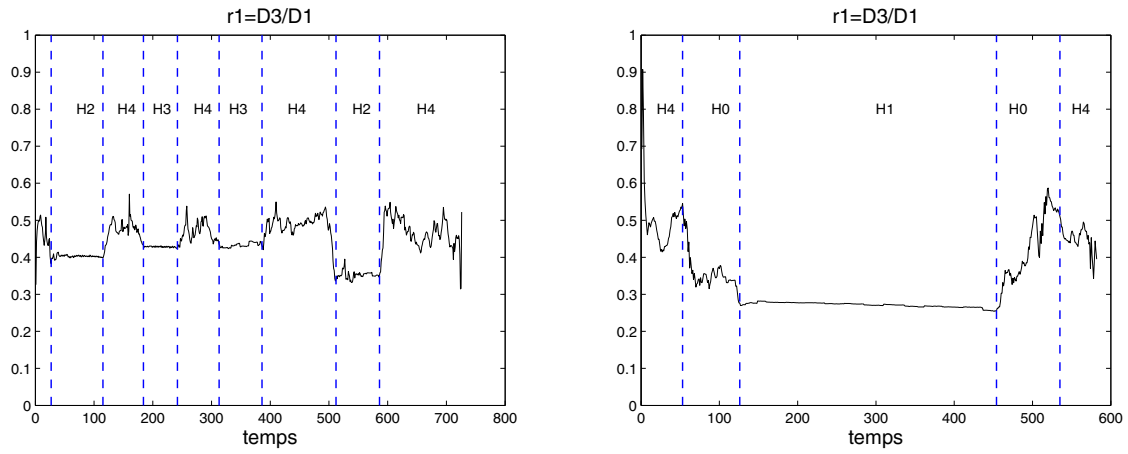


FIG. 3.27 – Variations de  $r_1$  suivant deux scénarios.

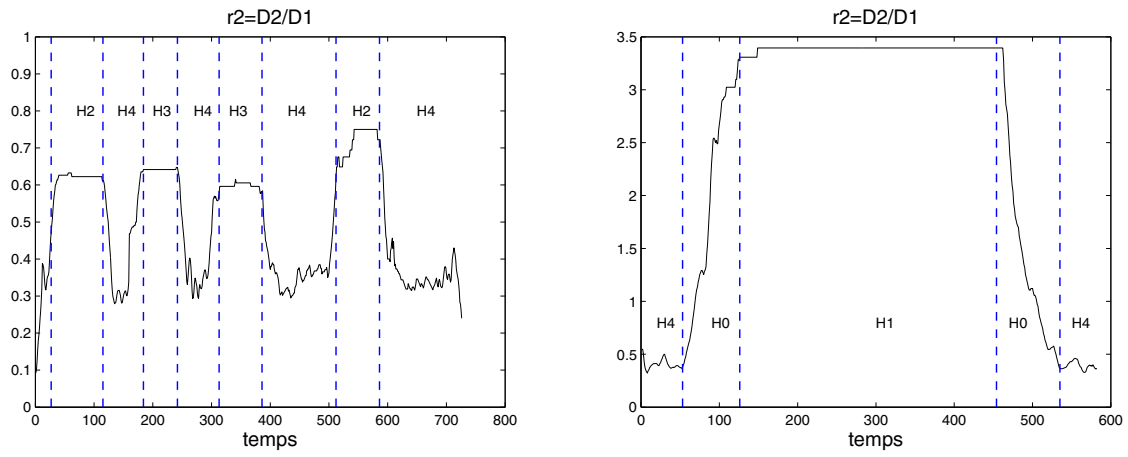


FIG. 3.28 – Variations de  $r_2$  suivant deux scénarios.

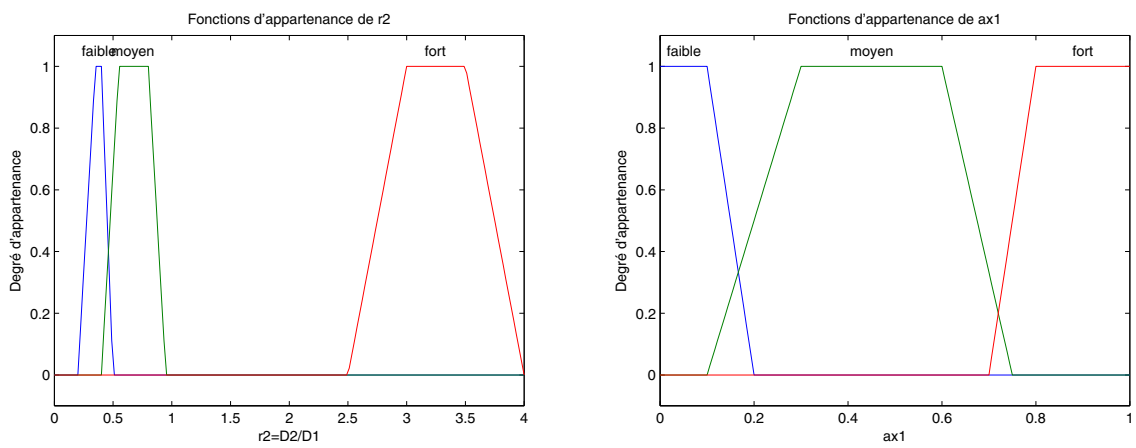


FIG. 3.29 – Fonctions d'appartenance des entrées  $r_2$  et  $ax_1$ .

$$\text{SI } (x \in A) \text{ ET } (y \in B) \text{ ALORS } (z \in C)$$

où  $A$ ,  $B$  et  $C$  sont des sous-ensembles flous.

Il convient de déterminer l'ensemble de ces raisonnements, également appelés règles d'inférence.

Pour notre application, 8 règles ont été définies tenant compte de l'ensemble des entrées et de l'univers de sortie. Le tableau 3.5 résume l'ensemble de ces règles.

N°		Entrées					Sortie
		$r_1$	$r_2$	$ax_1$	$ax_2$		
1	SI			fort	faible	ALORS	couché
2	SI	moyen faible		moyen	fort	ALORS	couché
3	SI	moyen fort		moyen	fort	ALORS	assis
4	SI		moyen	faible	fort	ALORS	accroupi
5	SI		faible	faible	fort	ALORS	debout
6	SI	faible	fort			ALORS	couché
7	SI	fort	faible			ALORS	debout
8	SI	moyen fort	moyen			ALORS	assis

TAB. 3.5 – Règles définies pour le raisonnement de l'algorithme de logique floue.

### **Remarque 3.25**

- Comme pour la détermination des fonctions d'appartenance, la définition des règles de combinaison est réalisée à partir de constatations expérimentales au travers des courbes obtenues lors de la campagne de tests.
- Les autres configurations correspondent à une posture indéterminée ou à une situation impossible.

### **(c) Les opérateurs d'implication, de conjonction et d'aggrégation**

L'étape suivante consiste à opérer un choix quant aux opérateurs d'implication et de conjonction traduisant respectivement le ALORS et le ET. Nous avons opté pour l'implication floue de Mamdani [Mam77].

Le degré de l'implication floue pour une règle du type :

$$\text{SI } (x \in A_i) \text{ ET } (y \in B_i) \text{ ALORS } (z \in C_i)$$

est défini par :

$$f_{S_i}(z) = \min(\min(f_{A_i}(x), f_{B_i}(y)), f_{C_i}(z)). \quad (3.6)$$

L'opérateur d'implication et l'opérateur de conjonction correspondent donc à l'opérateur min.

Exemple : Soit le vecteur d'entrée  $V = (r_1, r_2, ax_1, ax_2)^T = (0.5, 2.9, 0.9, 0.23)^T$ . D'après la règle N°1, on obtient :

Si ( $ax_1$  est fort) ET ( $ax_2$  est faible) ALORS (posture est couchée)

$ax_1$  est fort avec un degré de 1 (figure 3.30 (a)) et  $ax_2$  est faible avec un degré de 0.33 (figure 3.30 (b)). On déduit la fonction d'appartenance des postures solutions de l'implication floue sur la figure 3.30 (c) ( $\min(1; 0.33) = 0.33$ ).

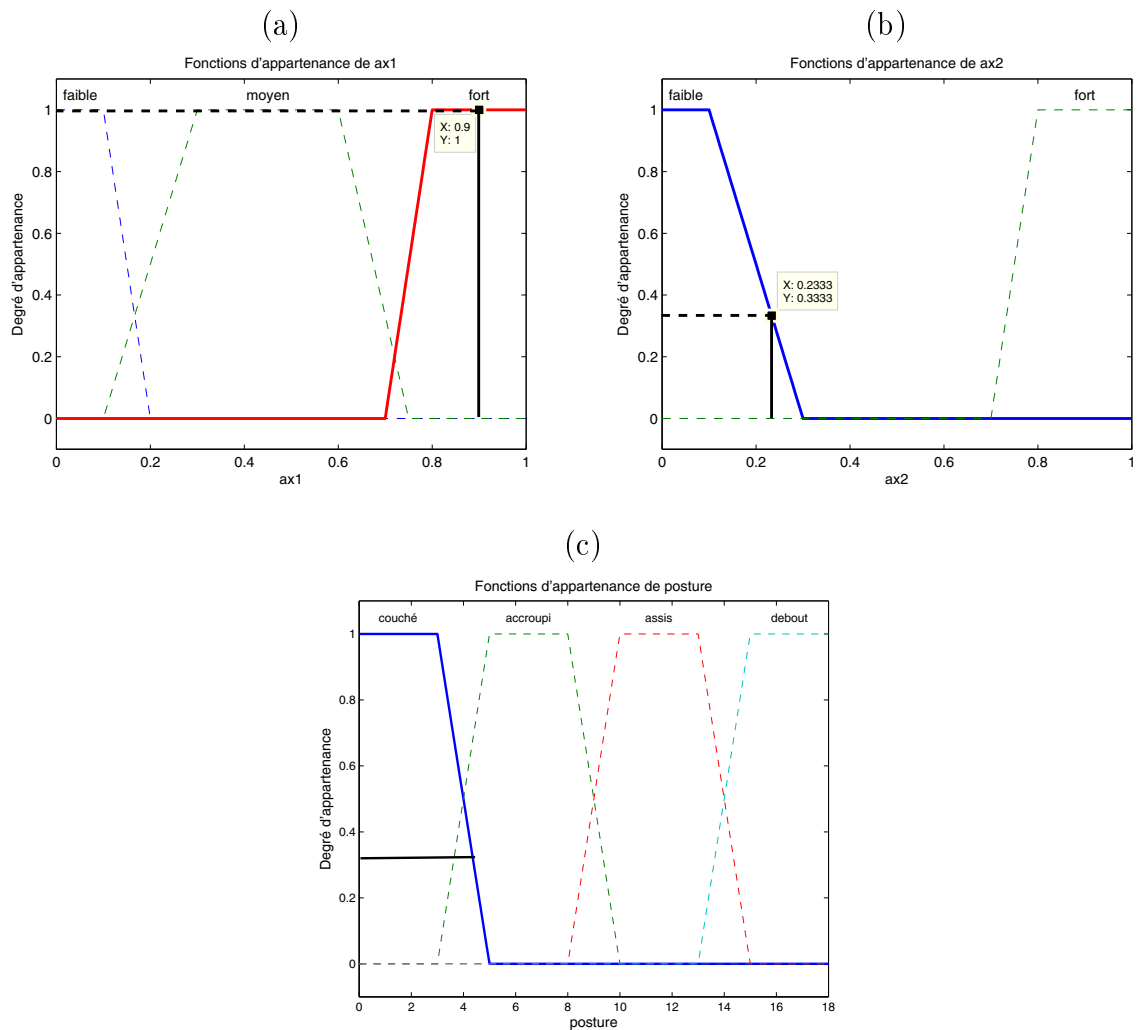


FIG. 3.30 – Illustration de la règle N°1.

Un travail similaire est effectué pour chacune des 7 autres règles avec le même ensemble de valeurs d'entrée. A partir de ces fonctions d'appartenance, on détermine la fonction d'appartenance globale du système en agrégeant les résultats à l'aide de l'opérateur max :

$$f_S(z) = \max(f_{S_1}(z), f_{S_2}(y), \dots, f_{S_8}(z)). \quad (3.7)$$

Pour l'ensemble d'entrées  $(0.5, 2.9, 0.9, 0.23)^T$ , la fonction d'appartenance solution est illustrée à la figure 3.31.

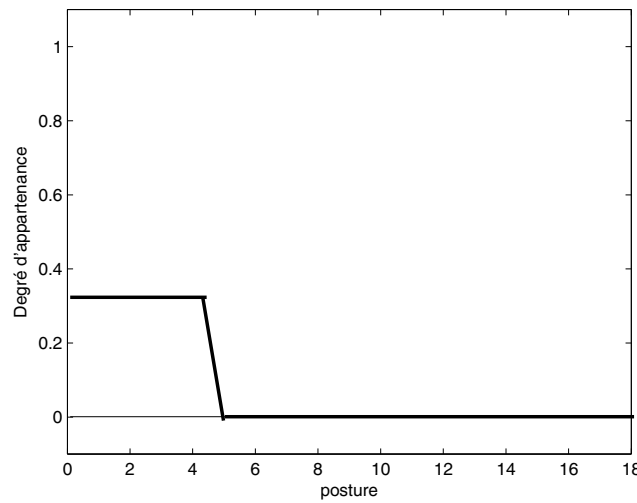


FIG. 3.31 – Fonction d'appartenance globale du système.

#### (d) La défuzzification

La dernière étape, appelée défuzzification, permet d'interpréter cette fonction d'appartenance solution. On distingue différentes méthodes :

- la méthode du maximum où la solution correspond à l'abscisse du maximum de la fonction d'appartenance globale. Dans le cas où plusieurs points conviennent, trois alternatives sont possibles : (1) Mean Of Max (MOM), moyenne des maxima, (2) Smallest of Max (SOM), minimum des abscisses des maxima, (3) Largest Of Max (LOM), maximum des abscisses des maxima ;
- la méthode du centre de gravité : elle offre de bonnes performances en terme de stabilité face aux variations des entrées, ce qui explique qu'elle est largement utilisée. Cependant elle nécessite une puissance de calcul plus importante.

**Remarque 3.26** Nous avons choisi la méthode de la moyenne des maxima car elle ne nécessite pas une grande puissance de calcul et est implémentable en temps réel.

### 3.3.3.2 Résultats expérimentaux

Un ensemble de 15 séquences vidéos a été réalisé, soit un total d'environ 7700 images (figure 3.32). Chacune des séquences a permis de représenter différents scénarios que l'on peut retrouver dans la vie courante ainsi que des scénarios de chute. L'objectif est de tester la robustesse et les performances du dispositif de reconnaissance de posture.

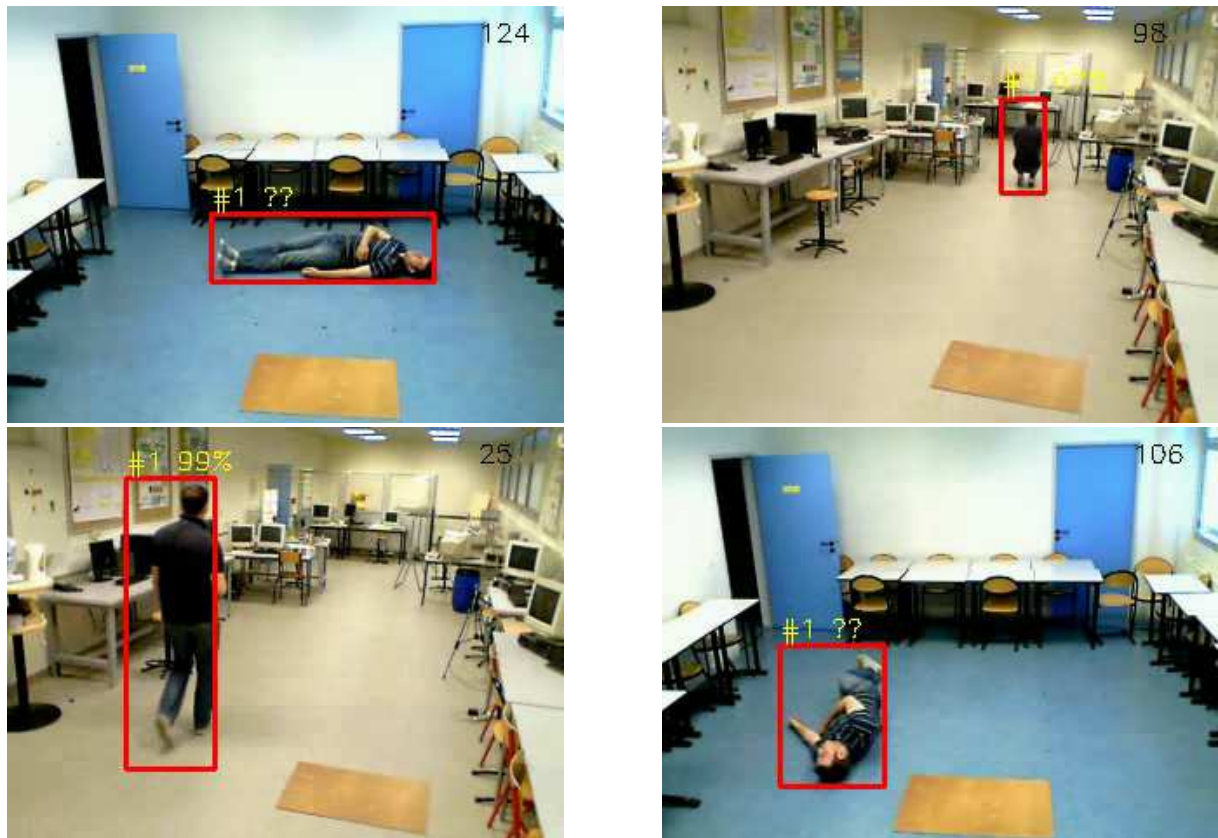


FIG. 3.32 – Exemple de postures utilisées pour les tests.

Les taux de reconnaissance obtenus pour chaque posture, à partir de ces 15 séquences, sont donnés dans le tableau 3.6. La première ligne représente la posture réelle de la personne et les colonnes représentent la posture reconnue par le système de fusion.

Syst/Réel	$H_1$	$H_2$	$H_3$	$H_4$	$H_0$
$H_0$	1,2%	0,0%	0,0%	1,3%	8,2%
$H_1$	<b>83,1%</b>	0,0%	0,3%	0,4%	13,7%
$H_2$	6,5%	<b>83,0%</b>	4,9%	4,5%	25,9%
$H_3$	9,1%	14,9%	<b>94,8%</b>	1,8%	32,8%
$H_4$	0,0%	2,1%	0,0%	<b>92,0%</b>	19,5%

TAB. 3.6 – Matrice de confusion du système de reconnaissance de posture.

Le taux de reconnaissance moyen du système est de : 88,2%. Les différentes postures sont dans l'ensemble bien reconnues : des confusions existent entre les positions *accroupi* et *assis*. Les mauvaises reconnaissances du système peuvent s'expliquer par :

- les variations lumineuses et les zones d'ombres qui ont parfois un impact sur la taille de la boîte englobante ;
- les fonctions d'appartenance et des règles qui nécessiteraient quelques ajustements.

**Remarque 3.27** *Si la personne chute dans l'axe de la caméra, les données extraites de la boîte englobante ne permettent pas toujours de discriminer la posture.*

### 3.3.4 Fusion pour la prise de décision

Grâce aux trois modules de fusion précédents les registres relatifs à la détection, la localisation, la posture et le nombre de personnes ont été renseignés. A ces registres s'ajoutent des informations relatives à l'environnement extérieur telles la météo et l'heure. Différents niveaux de fusion ont été réalisés : une fusion de décision pour l'information de présence humaine, une fusion de primitives pour la localisation dans l'image et dans le réel et une fusion de primitives pour la détermination de posture. Des fusions de niveau 0 et 1 sont également opérées notamment lors du fonctionnement du VRHE. La dernière étape consiste donc à utiliser l'ensemble de ces registres pour établir un diagnostic en fonction de l'application étudiée.

### (a) La gestion énergétique d'habitat intelligent

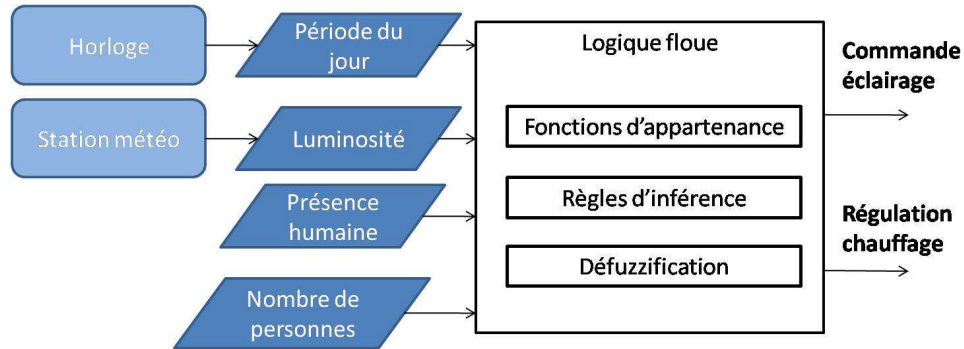


FIG. 3.33 – Prise de décision pour la gestion énergétique dans l'habitat.

L'objectif pour des applications d'habitat intelligent est double. D'une part, on cherche à assurer le confort de la personne au sein de chacune des pièces en automatisant la gestion de l'éclairage et du chauffage. D'autre part, on cherche à diminuer la consommation énergétique globale de l'habitat.

Si la présence humaine est détectée, il faut tenir compte des données de l'environnement extérieur. En effet, pourquoi éclairer une pièce si une personne est détectée et que la luminosité extérieure est suffisante pour assurer le confort visuel de la personne. Une prise en compte d'informations issues de l'environnement extérieur est nécessaire, afin de reproduire le comportement qu'aurait un humain si il devait manuellement allumer le système d'éclairage. L'objectif est de réduire efficacement la consommation et non pas de consommer inutilement de l'énergie.

La période du jour et la luminosité sont cependant des notions imprécises et interprétées de manières différentes suivant les personnes (elles seront simulées pour les exemples d'application). Afin de pallier cette difficulté, nous définissons ces deux informations à l'aide de sous-ensembles flous. Pour la période de la journée nous considérons trois sous-ensembles : la journée, la soirée et la nuit. Deux sous-ensembles sont également pris en compte pour la luminosité extérieure : sombre et clair. Enfin, la troisième entrée concernant la présence est composée de deux sous-ensembles : absence et présence humaine. Le système ainsi défini permet de gérer en sortie l'éclairage intérieur de la pièce, que nous divisons en deux sous-ensembles : OFF et ON. Les fonctions d'appartenance des trois entrées (présence, luminosité, période du jour) et de la sortie (éclairage) sont représentées par la figure 3.34.

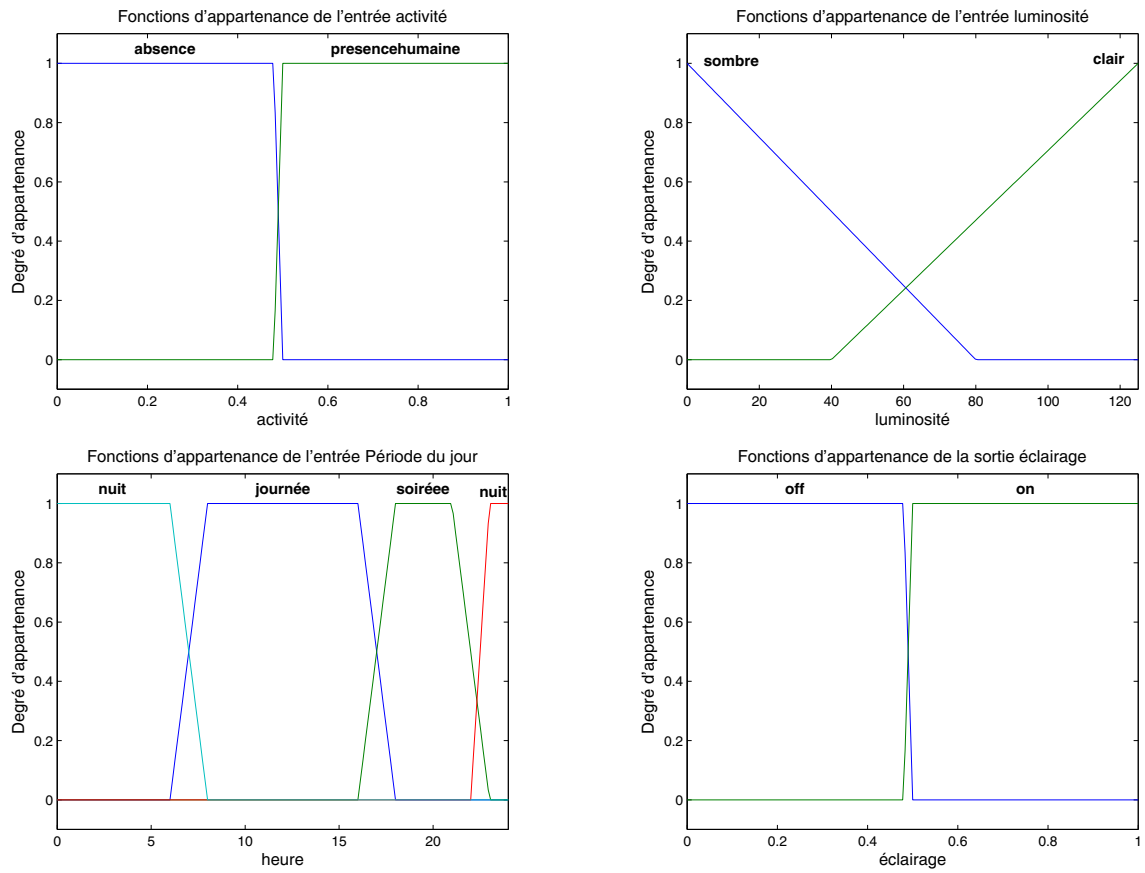


FIG. 3.34 – Fonctions d'appartenance du système d'éclairage.

Des règles portant sur ces trois entrées sont ensuite définies (tableau 3.7).

N°		Entrées				Sortie
		Activité	Luminosité	Période jour		
1	SI	absence			ALORS	OFF
2	SI	présence humaine		nuit	ALORS	ON
3	SI	présence humaine	normale	journée	ALORS	OFF
4	SI	présence humaine	sombre	journée	ALORS	ON
5	SI	présence humaine	normale	soirée	ALORS	OFF
6	SI	présence humaine	sombre	soirée	ALORS	OFF

TAB. 3.7 – Règles définies pour le raisonnement de l'algorithme de logique floue.

Le dispositif adapte donc son comportement en fonction de la période du jour et de la luminosité afin de couvrir le maximum de situations possibles.



**Remarque 3.28** Un raisonnement similaire peut être appliqué à la gestion du chauffage avec en entrée la présence humaine, le nombre de personnes, la température extérieure et la période du jour. Une régulation de la température intérieure est alors envisageable.

## (b) L'aide au maintien à domicile

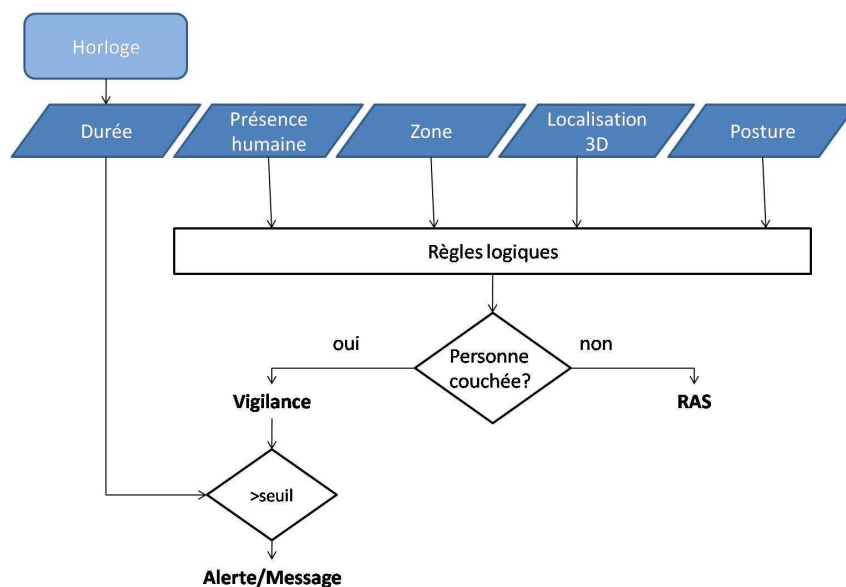


FIG. 3.35 – Prise de décision pour le maintien à domicile.

L'un des principaux risques, pour les personnes âgées, au sein de l'habitat, est la chute. Pour permettre la détection de chute, le dispositif CAPTHOM doit tout d'abord déterminer la posture de la personne. Cette phase de détermination de posture est effectuée à l'aide du module présenté au paragraphe (3.3.3). Cependant la posture *couché* ne signifie pas forcément que la personne a chuté. Elle peut très bien être couchée dans son lit, ce qui constitue une situation normale de la vie courante. Afin d'évaluer correctement la situation, une donnée supplémentaire doit être prise en compte : la localisation de la personne issue du module présenté au paragraphe (3.3.2). Le fonctionnement d'un système combinant l'information *posture* et l'information *localisation* nécessite des connaissances a priori sur la localisation du lit dans l'environnement réel et/ou dans l'image. Des réglages du dispositif doivent être opérés au préalable lors de l'installation du CAPTHOM. Cela suppose également que la disposition de la pièce ne doit pas changer afin que le dispositif continue à délivrer une information pertinente. Cette hypothèse se justifie par le fait que, chez les personnes âgées, il est rare que les meubles soient déplacés. Enfin, lorsque la posture *couché* est déterminée, un compteur est mis en marche afin de connaître la durée pendant laquelle la personne reste couchée. Le calcul de cette durée permet de diminuer le

taux de fausses détections, en fixant un seuil temporel au delà duquel la décision de posture *couché* est confirmée. Cela permet entre autre de ne pas prendre en compte les brèves variations des entrées du dispositif de détermination de posture qui pourrait conduire à la décision *personne couchée*.

Afin d'illustrer les décisions du dispositif CAPTHOM en fonction des informations délivrées par les capteurs, nous avons développé une interface qui permet de visualiser, à partir de scénarios représentatifs de situations courantes ou à risques, le comportement du dispositif et les décisions finales que ce soit pour une application de maintien à domicile ou de gestion énergétique. Un exemple de l'interface développée est présenté figure 3.36. Deux parties composent cette interface. La partie supérieure, qui a terme devrait disparaître, représente les données issues des capteurs (vidéo, station météo, horloge,...). Elle n'est présente que dans un but démonstratif. La partie inférieure permet de visualiser l'évolution de la valeur des différents registres et de la décision pendant le déroulement du scénario.

**Remarque 3.29** *L'ensemble des simulations a été réalisé hors ligne à partir de données récupérées via le traitement d'images et le logiciel de simulation SNOOPS.*

Différentes décisions prises par le système sont illustrées de la figure 3.36 à 3.40.



FIG. 3.36 – Situation de vigilance.

La figure 3.36 illustre l'envoi d'une alerte *Vigilance* étant donné que la personne a été détectée comme étant couchée dans une zone où cette posture est inhabituelle. Le message d'alerte *Chute* sera envoyé si la personne reste allongée pendant plus de 10 secondes.

Une situation de chute est présentée à la figure 3.37. La personne est en effet en position couchée sur le sol depuis plus de 10 secondes. Un message d'alerte signalant la chute est donc envoyé à la famille et au personnel soignant. Bien que la personne soit allongé dans l'axe de la caméra, situation posant des problèmes de reconnaissance, la posture est bien identifiée.



FIG. 3.37 – Scénario de chute.



FIG. 3.38 – Non déclenchement du dispositif d'éclairage de la pièce.

Sur la figure 3.38, l'entrée de la personne dans la pièce n'a pas déclenché l'éclairage dans la mesure où le scénario se déroule en journée et que la luminosité est normale. La luminosité est donc suffisante pour assurer le confort visuel de la personne.



FIG. 3.39 – Déclenchement du dispositif d'éclairage en fin de journée.

En revanche, si la même scène se déroule en fin de journée et que le temps est couvert (fin de journée d'hiver par exemple), le dispositif déclenche l'éclairage de la pièce (figure 3.39). On constate, dans ces deux derniers scénarios, que la position debout est bien identifiée.

Enfin, une situation d'absence est illustrée à la figure 3.40.



FIG. 3.40 – Absence de présence humaine.

## 3.4 Conclusion

Nous avons montré l'intérêt de fusionner les différentes données disponibles. Nous avons développé quatre modules de fusion avec des niveaux de fusion différents :

- le module de détection de présence humaine est basé sur une fusion de décision de par la nature hétérogène des données ;
- le module de localisation est basé sur une fusion de capteurs pour la localisation par zone et une fusion de primitives pour la localisation précise ;
- le module de détermination de la posture est basé sur une fusion de primitives issues du traitement de l'image. Aucun capteur n'est embarqué sur la personne.
- le module de prise de décision est basé sur la fusion de décisions.

Différentes méthodes (logique floue, théorie des ensembles) ont été utilisées.

Ce travail s'inscrivant dans un projet industriel, une interface graphique permettant de visualiser l'information des différents registres et la décision finale en fonction de l'application considérée, a été mise au point. Des résultats expérimentaux en laboratoire illustrent les potentialités de l'approche développée. Nous n'avons hélas pas pu tester les algorithmes sur la plateforme du GIS Madonah. En effet, pour des raisons éthiques, l'appartement témoin de la maison de retraite n'est actuellement pas équipé de caméra.







# Conclusion

Dans ce travail, nous avons proposé un dispositif de fusion de données multi-capteurs pour l'habitat intelligent. Nous nous sommes plus particulièrement intéressés au maintien à domicile de personnes âgées et à la gestion énergétique optimisée de bâtiments.

Nous avons montré que la détection de présence humaine était une problématique complexe, impliquant la détection non seulement des mouvements mais également de l'absence de mouvement de la personne. Trois types de capteurs pour le projet CAPTHOM ont été retenus : les détecteurs IRP (pour la détection de mouvement), les thermopiles (pour la détection de présence) et la caméra (pour l'identification entre autres). Ces informations, redondantes et complémentaires, sont alors fusionnées pour statuer sur la présence d'une personne dans un environnement surveillé, même en présence de perturbations.

Une fois la présence d'une personne détectée, il convient ensuite de la localiser. Nous avons montré que l'estimation visuelle sur horizon glissant apporte une solution à la localisation 3D d'une personne à partir d'informations issues de l'image. Le problème de la localisation 3D est reformulé en un problème d'optimisation non linéaire sous contraintes dans le plan image. L'estimation de la position 3D est assurée par un modèle global, déterminé à partir des modèles du déplacement humain et du système de vision. Véritable pièce maîtresse, il assure l'estimation de la position de la personne dans l'image sur un horizon et confère à l'approche une grande flexibilité. Un modèle de locomotion non holonome (modèle non linéaire) représentant plus fidèlement le déplacement humain a été utilisé.

Nous avons également étudié la détermination de la posture pour des applications de maintien à domicile de personnes âgées. Un dispositif, sans capteur embarqué sur la personne, a été développé. Quatre paramètres définis dans l'image constituent les entrées du module de fusion basé sur la logique floue. Le taux de reconnaissance est satisfaisant et peut être amélioré grâce à des images de meilleure qualité et des règles d'inférence affinées.

Enfin, un module de prise de décision permet de fusionner les diverses informations issues soit directement de capteurs, soit des autres modules de fusion (détection, localisation, détermination de la posture).

Pour répondre aux attentes des industriels, partenaires du projet, nous avons mis au point une interface graphique qui offre la possibilité de visualiser l'information des différents registres et la décision finale en fonction de l'application considérée.

Si les expérimentations ont permis de montrer l'efficacité du dispositif proposé, ce dernier n'a pas été validé en conditions réelles d'habitat intelligent. Nous envisageons de le faire dans l'appartement témoin du GIS MADONAH. Il nous faudra tout d'abord convaincre le personnel et les futurs occupants de la maison de retraite de l'intérêt avéré de l'image pour la surveillance à domicile. L'image est un signal, non visualisé à l'extérieur et supprimé après traitement. L'image ne porte donc pas atteinte à la vie privée des personnes.

On portera une attention particulière aux difficultés relatives à la synchronisation des données.

De plus, ces expérimentations nous permettraient d'affiner les règles d'inférence des différents modules et de connaître l'incertitude du dispositif de mesures (fiabilité des capteurs). Cette connaissance de l'incertitude (modélisation) pourrait alors être prise en compte en utilisant la théorie des possibilités au lieu de la logique floue.





# Publications de l'auteur

## Conférences internationales avec actes et comité de lecture

1. Brulin D., Courtial E., Becis-Aubry Y., Touré Y. Suivi de personnes par fusion de données et filtrage de Kalman. *In Proc. of Conférence Internationale Francophone d'Automatique (CIFA)*, Bucarest, Roumanie, September 2008.
2. Brulin D., Courtial E., Allibert G. Visual Receding Horizon Estimation for human presence detection. *In Proc. of IEEE International Conference on Robotics and Automation (ICRA) – Workshop on People Detection and Tracking*, Kobe, Japon, May 2009.
3. Brulin D., Courtial E., Allibert G. Position Estimation and Fall Detection using Visual Receding Horizon Estimation. *In Proc. of the 48th Conference on Decision and Control (CDC)*, Shangai, China, December 2009.
4. Brulin D., Courtial E. Multi-sensors data fusion system for fall detection. *In Proc. of the 10th IEEE International conference on Information Technology and Application in Biomedecine*, Corfu, Greece, November, 2010.

## Conférences nationales avec actes

1. Brulin D., Courtial E., Becis-Aubry Y., Touré Y. Suivi de cibles par fusion de données et filtrage de Kalman. *In Proc. of Sixième Colloque Capteurs*, Bourges, France, Mars 2008.



# Bibliographie

- [And08] Andriluka M., Roth S., Schiele B. People-tracking-by-detection and people-detection-by-tracking. In *Proc. of the Conference on Computer Vision and Pattern Recognition*, Anchorage, Alaska, June 2008.
- [Are06a] Arechavaleta G., Laumond J.-P., Hicheur H., Berthoz A. Optimizing principles underlying the shape of trajectories in goal oriented locomotion for humans. In *Proc. of the 6th IEEE-RAS International Conference on Humanoid Robots*, pages 131–136, Genova, Italy, December 2006.
- [Are06b] Arechavaleta G., Laumond J.P., Hicheur H., Berthoz A. The nonholonomic nature of human locomotion : a modeling study. In *Proc. of the International Conference on Biomedical Robotics and Biomechatronics*, Pisa, Italy, February 2006.
- [Bai98] Bai M., Zhou D.H., Schwarz H. Adaptative augmented state feedback control for an experimental planar two-link flexible manipulator. *IEEE Transactions on Robotics and Automation*, 14 :894–901, 1998.
- [Baj09] Bajracharya M., Moghaddam B., Howard A., Brennan S., Matthies L.H. Result from a real-time stereo-based pedestrian detection system on a moving vehicle. In *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)– Workshop on People Detection and Tracking*, Kobe, Japan, May 2009.
- [Bal04] Baldinger J.L., Boudy J., Dorizzi B., Levrey J.P., Andreo R., Perpère C., Devault F., Rocaries F., Deitrich C., Lacombe A. Tele-surveillance System for Patient at Home : The MEDEVILLE system. In *Proc. of the International Conference on Computers Helping People with Special Needs*, Paris, France, July 2004.
- [Ban03] Banerjee S., Steenkeste F., Couturier P., Debray M., Franco A. Telesurveillance of elderly patients by use of passive infra-red sensors in a smart room. *Journal of Telemedicine and Telecare*, 9(1) :23–29, 2003.
- [Bar05] Barralon P., Noury N., Vuillerme N. High level information extracted from a kinematic sensor to monitor physical activity. In *Proc. of the 27th Annual*



- International Conference IEEE Engineering in Medicine and Biology Society"*, pages 1703–1706, Shangai, China, September 2005.
- [Bel00] Belongie S., Malik J., Puzicha J. Shape Context : A new descriptor for shape matching and object recognition. In *Proc. of the Neural Information Processing System*, pages 831–837, Denver, USA, June 2000.
  - [Bel02] Bellot D., Boyer A., Charpillat F. A new definition of qualified gain in a data fusion process : application to telemedicine. In *Proc. of Fusion*, Annapolis, Maryland, USA, July 2002.
  - [Bel08] Belconde A., Kratz F. Problèmes liés à la détection de présence humaine et modélisation. In *Proc. of 6ème Colloque Capteurs*, Bourges, France, March 2008.
  - [Ben08] Benezeth Y., Emile B., Laurent H., Rosenberger C. A Real Time Human Detection System Based on Far Infrared Vision. In *Proc. of the 3rd International Conference on Image and Signal Processing (ICISP)*, pages 76–84, Cherbourg-Octeville, France, July 2008.
  - [Ben09a] Benezeth Y. *Détection de la présence humaine par vision*. PhD thesis, Université d'Orléans, Octobre 2009.
  - [Ben09b] Benezeth Y., Emile B., Laurent H., Rosenberger C. Détection de la présence humaine et caractérisation de l'activité. In *Proc. of the XXIIe colloque GRETSI (traitement du signal et des images)*, Dijon, France, September 2009.
  - [Bey99] Beymer D., Konolige K. Real-time tracking of multiple people using continuous detection. In *Proc. of IEEE International Conference on Computer Vision (ICCV) – Frame-rate Workshop*, Corfu, Greece, September 1999.
  - [Bla03] Blasch E.P., Plano S. Level 5 : User refinement to aid the fusion process. In *Proc. of Multisensor, Multisource Information Fusion : Architectures, Algorithms and Applications*, volume 5099, pages 288–297, Orlando, Florida, USA, April 2003.
  - [Blo03] Bloch I. *Fusion d'Informations en Traitement du Signal et des Images*. Hermes Sciences Publications, 2003. ISBN 2-7462-0628-5.
  - [Bon08a] Bonhomme S. *Méthodologie et Outils pour la Conception d'un Habitat Intelligent*. PhD thesis, Institut National Polytechnique de Toulouse, Mai 2008.
  - [Bon08b] Bonhomme S., Campo E., Estève D., Guennec J. PROSAFE-extended, a telemedicine platform to contribute to medical diagnosis. *Journal of Telemedicine and Telecare*, 14(3) :116–119, 2008.
  - [Bou06] Boudy J., Baldinger J.L., Delavault F., Muller M., Farin I., Andreao R., Torres-Müller S., Serra A., Gaïti D., Rocaries F., Lacombe A., Steenkeste F., Schaff

- M., Baer M., Ozguler A., Vaysse S. Telemedecine for elderly patient at home : the Telepat project. In *Proc. of the International Conference On Smart Homes and Health Telematics*, Belfast, Northern Ireland UK, June 2006.
- [Bow04] Bowman C.L. The Dual Node Network (DNN) data fusion and ressource management architecture. In *Proc. of the AIAA Intelligent Systems Conference*, Chicago, Illinois, USA, September 2004.
- [Bro86] Broida T., Chellappa R. Estimation of object motion parameters from noisy images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1 :90–99, 1986.
- [Bru09a] Brulin D., Courtial E., Allibert G. Position Estimation and Fall Detection using Visual Receding Horizon Estimation. In *Proc. of the 48th Conference on Decision and Control (CDC)*, Shanghai, China, December 2009.
- [Bru09b] Brulin D., Courtial E., Allibert G. Visual Receding Horizon Estimation for human presence detection. In *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)– Workshop on People Detection and Tracking*, Kobe, Japan, May 2009.
- [Cas09] Castaño B., R-Moreno M.D. An Integrated AI and RFID System for People Detection and Orientation. *Journal on Systemics, Cybernetics and Informatics*, 7(1) :49–55, 2009.
- [Cha03] Chan M., Campo E., Estève D. PROSAFE, a multisensory remote monitoring system for the elderly or the handicapped. In *Proc. of the 1st International Conference on Smart Homes and Health Telematics*, pages 89–95, Paris, France, September 2003.
- [Cha08] Chan M., Estève D., Escriba C., Campo E. A review of smart homes-Present state and future challenges. *Computer Methods and Programs in Biomedicine*, 91(1) :55–81, 2008.
- [Col01] Collins, R.T. Lipton, A.J. Fujiyoshi, H. Kanade, T. Algorithms for Cooperative Multisensor Surveillance. *Proceedings of the IEEE*, 89(10) :1456 – 1477, October 2001.
- [Coo03] Cook D.J., Youngblood M., Heierman E.O. III, Gopalratnam K., Rao S., Litvin A., Khawaja F. MavHome : An Agent-Based Smart Home. In *Proc. of the First IEEE International Conference on Pervasive Computing and Communications*, pages 521–524, Fort Worth, Texas, USA, March 2003.
- [Cou96] Courtial E. *Commande prédictive et estimation d’état de systèmes non linéaires*. PhD thesis, Université Claude Bernard, Lyon 1, Avril 1996.

- [Cro93] Crowley J.L., Demazeau Y. Principles and Techniques for Sensor Data Fusion. *Signal Processing*, 32(1-2) :5–27, 1993.
- [Dal05] Dalal N., Triggs B. Histograms of Oriented Gradients for Human Detection. In *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 886–893, San Diego, USA, June 2005.
- [Dal06] Dalal N., Triggs B., Schmid C. Human Detection Using Oriented Histograms of Flow and Appearance. In *Proc. of the European Conference on Computer Vision*, pages 428–441, Graz, Austria, May 2006.
- [Dav07] David P., Idasiak V., Kratz F. A sensor placement approach for the monitoring of indoor scenes. In *Proc. of the European Conference on Smart Sensing and Context*, volume 4793, Kendal, England, October 2007.
- [Dav09] David P. *Contribution à l'analyse de sûreté de fonctionnement des systèmes complexes en phase de conception : application à l'évaluation de réseaux de capteurs de présence humaine*. PhD thesis, Université d'Orléans, Novembre 2009.
- [Dem67] Dempster A.P. Upper and lower probabilities induced by a multivalued mapping. *The Annals of Mathematical Statistics*, 38(2) :325–339, 1967.
- [Dem68] Dempster A.P. A Generalization of Bayesian Inference. *Journal of the Royal Statistical Society Series B*, 30(2) :205–247, 1968.
- [Dem02] Demongeot J., Virone G., Duchêne F., Benchetrit G., Hervé T., Noury N., Rialle V. Multi-sensors acquisition, data fusion, knowledge mining and alarm triggering in health smart homes for elderly people. *Comptes Rendus Biologies*, 325(6) :673–682, 2002.
- [Duc04] Duchêne F. *Fusion de données multicapteurs pour un système de télésurveillance médicale de personnes à domicile*. PhD thesis, Université Joseph Fourier, Lyon 1, Octobre 2004.
- [Dud01] Duda R.O., Hart P.E., Stork D.G. *Pattern Classification*. Wiley-Interscience, 2nd edition, 2001.
- [Erm07] Ermanel C., Thélot B., Jouglu E., Pavillon G. Mortalité par accident de la vie courante en France métropolitaine, 2000-2004. Bulletin épidémiologique hebdomadaire n°37-38 de l'Institut de Veille Sanitaire, pages 318–322, Octobre 2007.
- [Esc05] Escriba C. *Conception, réalisation et caractérisation de capteurs infrarouges à thermopiles : application à la détection de présence passive dans l'habitat*. PhD thesis, Université Paul Sabatier Toulouse 3, Décembre 2005.

- [Ess09] Ess A., Schindler K., Leibe B., Van Gool L. Improved Multi-Person Tracking with Active Occlusion Handling. In *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)– Workshop on People Detection and Tracking*, Kobe, Japan, May 2009.
- [Fle08a] Fleury A., Noury N., Vacher M., Glasson H., Serignat J.-F. Sound and Speech Detection and Classification in a Health Smart Home. In *Proc. of the 30th IEEE EMBS Annual International Conference, "Personalized Healthcare through Technology"*, pages 4644–4647, Vancouver, Canada, August 2008.
- [Fle08b] Fleury A., Vacher M., Glasson H., Serignat J.F., Noury N. Data Fusion in Health Smart Home : Preliminary Individual Evaluation of Two Families of Sensors. In *Proc. of the 6th International Conference of the International Society for Gerontechnology (ISG'08)*, Pisa, Italy, May 2008.
- [Gir05] Girondel V., Caplier A., Bonnaud L. A belief theory-based static posture recognition system for real-time video surveillance applications. In *Proc. of the International Conference on Advanced Video and Signal Based Surveillance*, pages 10–15, Como, Italy, September 2005.
- [Gob06] Gobeau J.F. Détecteurs de mouvement à infrarouge passif. In *Proc. of Cinquièmes colloque Capteurs*, Bourges, France, October 2006.
- [Hal00] Hall M.J., Hall S.A., Tate T. Removing the HCI bottleneck : How the human computer interface (HCI) affects the performance of data fusion systems. In *Proc. of the MSS National Symposium on Sensor Data Fusion*, San Antonio, Texas, USA, June 2000.
- [Hal09] Hall D.L., Llinas J. Multisensor Data Fusion. In *Handbook of Multisensor Data Fusion : theory and practice*, chapter 1, pages 1–14. CRC Press, 2nd edition, 2009.
- [Har04] Hartley, R. I. and Zisserman, A. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN : 0521540518, second edition, 2004.
- [Has05] Haseltine E.L., Rawlings J.B. Critical Evaluation of Extended Kalman Filtering and Moving-Horizon Estimation. *Industrial and Engineering Chemistry Research*, 44(8) :2451–2460, 2005.
- [Hed07] Hedengren J.D., Allsford K.V., Ramlal J. Moving Horizon Estimation and Control for an Industrial Gas Phase Polymerization Reactor. In *Proc. of the American Control Conference (ACC)*, pages 1353–1358, New York City, USA, July 2007.

- [Ist06] Istrate D., Castelli E., Vacher M., Besacier L., Serignat J.F. Information extraction from sound for medical telemonitoring. *Transactions on Information Technology in Biomedicine*, 10 :264–274, 2006.
- [Joh09] Johnsen S., Tews A. Real-time object tracking and classification using a static camera. In *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)– Workshop on People Detection and Tracking*, Kobe, Japan, May 2009.
- [Kal60] Kalman R.E. A New Approach to Linear Filtering and Prediction Problems. *Transactions of the ASME–Journal of Basic Engineering*, 82(Series D) :35–45, 1960.
- [Kes09] Kessler O., White F. Data Fusion Perspective and its Role in Information Processing. In *Handbook of Multisensor Data Fusion : theory and practice*, chapter 2, pages 15–44. CRC Press, 2nd edition, 2009.
- [Kid99] Kidd C.D., Orr R., Abowd D., Atkesson C.G., Essa I.A., MacIntyre B., Mynatt E., Starner T.E., Newstetter W. The aware Home : A living Laboratory for ubiquitous computing research. In *Proc. of the 2nd International Workshop on cooperative Buildings*, Pittsburgh, USA, October 1999.
- [Kli95] Klir G.J., Yuan B. *Fuzzy Sets and Fuzzy Logic : Theory and Applications* . Prentice Hall, 1st edition, 1995.
- [Le 08] Le X.H.B. *Reconnaissance des comportements d’une personne âgée vivant seule dans un habitat intelligent pour la santé*. PhD thesis, Université Joseph Fourier Grenoble 1, Décembre 2008.
- [Lee90] Lee S., Kay Y. An accurate estimation of 3D position and orientation of a moving object for robot stereo vision : Kalman filter approach. In *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 414–419, Cincinnati, USA, May 1990.
- [Leg05] Legrand. Inone by legrand, 2005. [http ://www.inonebylegrand.fr/projet\\_0.php](http://www.inonebylegrand.fr/projet_0.php).
- [Lei05] Leibe B., Seemann E., Schiele B. Pedestrian Detection in Crowded Scenes. In *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 878–885, San Diego, USA, June 2005.
- [Li 09] Li Q., Stankovic J.A., Hanson M., Barth A., Lach J. Accurate, Fast Fall Detection Using Gyroscopes and Accelerometer-Derived Posture Information. In *Proc. of Body Sensor Network*, Berkeley ,USA, June 2009.
- [Lig09] Liggins M.E., Hall D.L., Llinas J. *Handbook of Multisensor Data Fusion : theory and practice*. The Electrical Engineering and Applied Signal Processing Series. CRC Press, 2nd edition, 2009.

- [Lip04] Lippiello V., Siciliano B., Villani L. Visual motion estimation of 3D objects : an adaptive extended Kalman filter approach. In *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, volume 1, pages 957–962, Sendai, Japan, September 2004.
- [Mah36] Mahalanobis P.C. On the generalised distance in statistics. In *Proc. of the National Institute of Science, India*, volume 2, pages 49–55, Calcutta, India, April 1936.
- [Mam77] Mamdani E.H. Applications of Fuzzy Set Theory to Control Systems : A Survey. In *Fuzzy Automata and Decision Processes*, pages 1–13. Gupta M.M., Saridis G.N. and Gaines B.R., eds., 1977.
- [Mas99] Massachusetts Institute of Technology. House\_n, 1999. [http://architecture.mit.edu/house\\_n](http://architecture.mit.edu/house_n).
- [Med08] Medjahed H., Istrate D., Boudy J., Baldinger J., Dorizzi B., Belfeki I., Martins V., Steenkeste F., Andreo R. A Multimodal Platform for Database Recording and Elderly People Monitoring. In *Proc. of The First International Conference on Biomedical Electronics and Devices*, pages 385–392, Funchal, Madeira, Portugal, January 2008.
- [Med09] Medjahed H., Istrate D., Boudy J., Dorizzi B. A fuzzy logic system for home elderly people monitoring (EMUTEM). In *Proc. of the 10th WSEAS International Conference on Fuzzy Systems*, pages 69–75, Prague, Czech Republic, March 2009.
- [Mik04] Mikolajczyk K., Schmid C., Zisserman A. Human Detection Based on a Probabilistic Assembly of Robust Part Detectors. In *Proc. of the European Conference on Computer Vision*, pages 69–82, Prague, Czech Republic, May 2004.
- [Nou00] Noury N., Hervé T., Rialle V., Virone G. Monitoring Behavior In Home Using a Smart Fall Sensor And Position Sensors. In *Proc. of the IEEE-EMBS Microtechnologies in Medicine and Biology*, pages 607–610, Lyon, France, October 2000.
- [Nou03] Noury N., Rialle V. Habitat intelligent pour la santé : systèmes et équipements. Les Techniques de l’Ingénieur, IN-9, 2003.
- [Nou04] Noury N., Virone G., Barralon P., Rialle V., Demongeot J. Maisons intelligentes pour personnes âgées : technologies de l’information intégrées au service des soins à domicile. *J3eA, Journal sur l’enseignement des sciences et technologies de l’information et des systèmes*, 3, 2004. Hors-Série 1.
- [Nou05] Noury N. AILISA : experimental platforms to evaluate remote care and assistive technologies in gerontology. In *Proc. of the 7th International Workshop on*

- enterprise networking and computing in Healthcare industry, Healthcom"*, pages 155–160, Busan, Korea, June 2005.
- [Pap98] Papageorgiou C., Oren M., Poggio T. A general framework for object detection. In *Proc. of the IEEE International Conference on Computer Vision*, pages 555–562, Bombay, India, January 1998.
- [Pap00] Papageorgiou C.P., Poggio T. A Trainable System for Object Detection. *International Journal of Computer Vision*, 38(1) :15–33, 2000.
- [Phi91] Philip J. Estimation of three dimensional motion of rigid objects from noisy observations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(1) :61–66, 1991.
- [Ram06] Ramanan D. Learning to parse images of articulated bodies. In *Proc. of the Advances in Neural Information Processing Systems (NIPS)*, Vancouver, Canada, December 2006.
- [Rei79] Reid D.B. An Algorithm for Tracking Multiple Targets. *Transactions on Automatic Control*, 24(6) :843–854, 1979.
- [Ria00] Rialle V., Lauvernay N., Piquard J.F., Franco A., Couturier P. Modélisation et expérimentation d’une chambre intelligente : les premiers pas d’une expérience hospitalière. *Télé médecine en gériatrie*, pages 105–127, 2000.
- [Ria01] Rialle V., Noury N., Fayn J., Chan M., Campo E., Bajolle L., Thomesse J.P. Health « smart » home information systems : concepts and illustrations. In *Proc. of the 3rd International Workshop on Enterprise Networking and Computing in Health Care Industry - Healthcom*, pages 99–103, L’Aquila, Italy, June 2001.
- [Ric93] Richardson S.J., Poulson D.F., Nicolle C. Supporting independent living through adaptable smart home (ASH) technologies. In *Proc. of the Human welfare and technologies : papers from the human service information technology applications (HUSITA) conference on information technology and the quality of life and services*, pages 87–95, Maastricht, Netherlands, June 1993.
- [Ruh07] Ruhm K.H. Sensor fusion and data fusion - Mapping and reconstruction. *Measurement : Journal of the International Measurement Confederation*, 40(2) :145–157, 2007.
- [Sat09] Satake J., Miura J. Robust Stereo-based person detection and tracking for a person following robot. In *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)– Workshop on People Detection and Tracking*, Kobe, Japan, May 2009.

- [Sch03] Schulz D., Burgard W., Fox D., Cremers A.B. People Tracking with a Mobile Robot Using Sample-Based Joint Probabilistic Data Association Filters. *International Journal of Robotics Research*, 32(2), 2003.
- [Sch09] Schiele B., Andriluka M., Majer N., Roth S., Wojek C. Visual People Detection — Different Models, Comparison and Discussion. In *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)– Workshop on People Detection and Tracking*, Kobe, Japan, May 2009.
- [See06] Seemann E., Leibe B., Schiele B. Multi-Aspect Detection of Articulated Objects. In *Proc. of the Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1582–1588, New York, USA, June 2006.
- [Sha76] Shafer G. *A Mathematical Theory of Evidence*. Princeton University Press, 1976.
- [Ste09] Steinberg A.N., Bowman C.L. Revision to the JDL Data Fusion Model. In *Handbook of Multisensor Data Fusion : theory and practice*, chapter 3, pages 45–68. CRC Press, 2nd edition, 2009.
- [Tyl96] Tyler M.L., Asano K., Morari M. A moving horizon estimation approach to robust fault detection. In *Proc. of the IMACS Multiconferences on Computational Engineering in Systems Applications*, pages 188–193, Lille, France, July 1996.
- [Vac09] Vacher M., Fleury A., Portet F., Serignat J-F., Noury N. Reconnaissance des sons et de la parole dans un Habitat Intelligent pour la Santé : expérimentations en situation non contrôlée. In *Proc. of the 13th Colloque du Groupe de Recherche et d’Etudes du Traitement du Signal et des Images (GRETSI)*, Dijon, France, September 2009.
- [Vio05] Viola P., Jones M.J., Snow D. Detecting pedestrians using patterns of motion and appearance. *International Journal of Computer Vision*, 63(2) :153–161, 2005.
- [Wan92] Wang J., Wilson W.J. 3D relative position and orientation estimation using Kalman filter for robot control. In *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 2638–2645, Nice, France, May 1992.
- [Wir00] Wira P., Urban J.P. A new adaptative Kalman filter applied to visual servoing tasks. In *Proc. of the 4th International conference on knowledge-based intelligent engineering systems and allied technologies*, pages 267–270, Brighton, England, August 2000.



- [Wu 07] Wu B., Nevatia R. Detection and Tracking of Multiple, Partially Occluded Humans by Bayesian Combination of Edgelet based Part Detectors. *International Journal of Computer Vision*, 75(2) :247–266, 2007.
- [Yeo08] Yeoh C.M., Tan H.Y., Kok C.K., Lee H.J., Lim H. e2Home : A Lightweight Smart Home Management System. In *Proc. of the 3rd International Conference on Convergence and Hybrid Information Technology*, pages 82–87, Busan, Republic of Korea, November 2008.
- [You08] Yous S., Laga H., Chihara K. People Detection and Tracking with World-Z Map from a Single Stereo Camera. In *Proc. of the International Workshop on Visual Surveillance*, Marseille, France, October 2008.
- [Zav08] Zavalla V.D., Laird C.D., Biegler L.T. A fast moving horizon estimation algorithm based on nonlinear programming sensitivity. *Journal of Process Control*, 18(9) :876–884, October 2008.

# Annexe A : Calcul de la matrice d'homographie

L'homographie est une application projective bijective qui permet d'effectuer la transformation entre deux plans [Har04]. La détermination de la matrice d'homographie  $M_{hom}$  nécessite l'utilisation d'un repère rectangulaire de dimensions connues afin de disposer d'au minimum de quatre points dont on connaît les coordonnées dans le repère environnement et les coordonnées pixelliques dans le plan image.

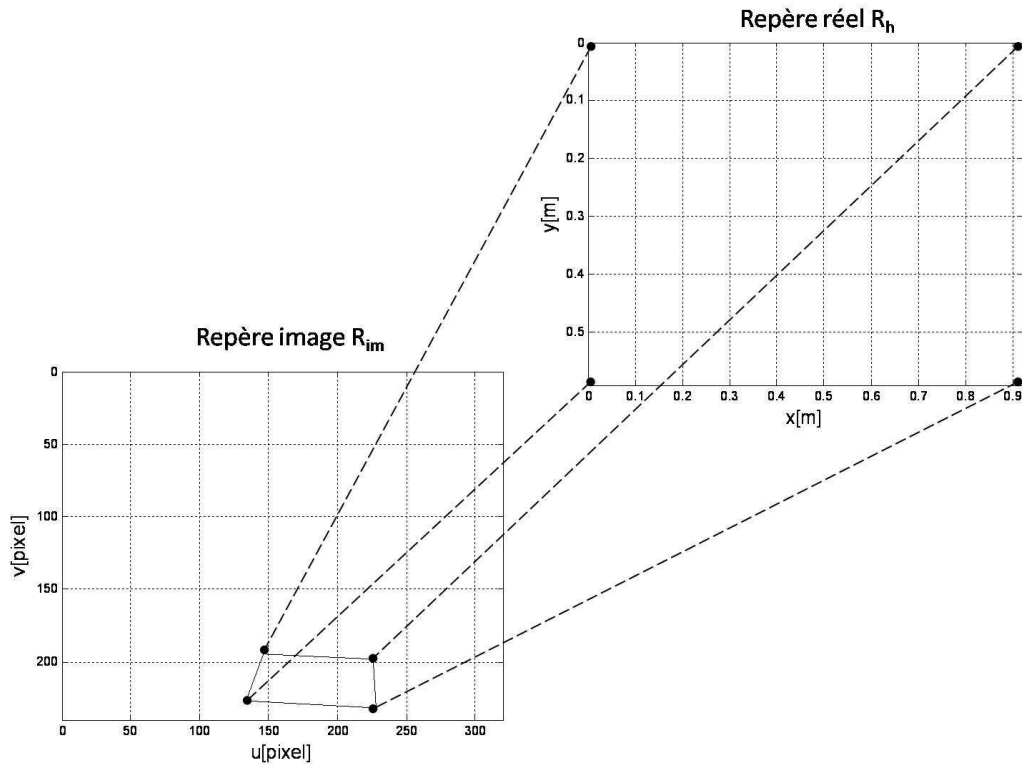


FIG. 3.41 – Correspondance entre les points pour le calcul d'homographie.

La matrice  $M_{hom}$  est calculée par une transformation linéaire directe. On définit pour

cela  $X'_i$  et  $X_i$ , les vecteurs contenant les coordonnées du sommet  $i$  (pour  $i$  allant de 1 à 4) respectivement sur la surface (repère  $R_h$ ) et dans le plan image  $R_{im}$  (figure 3.41). Afin de s'assurer que la matrice d'homographie ne dépende pas du système de coordonnées dans lequel les points sont exprimés, les deux ensembles de points sont normalisés séparément avant d'effectuer la transformation. On obtient alors pour chaque point les vecteurs normalisés  $\hat{X}'_i$  et  $\hat{X}_i$ .

La transformation linéaire directe se traduit alors par la relation suivante :

$$\hat{X}'_i = \hat{M}_{hom} \cdot \hat{X}_i \quad (3.8)$$

La notation  $\hat{M}_{hom}$  est utilisée pour rappeler qu'il sera nécessaire de dénormaliser à la fin de la transformation.

Cette transformation est valable pour chaque sommet  $i$  du repère. On peut donc écrire :

$$\begin{bmatrix} \hat{X}'_1 & \hat{X}'_2 & \hat{X}'_3 & \hat{X}'_4 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \cdot \begin{bmatrix} \hat{X}_1 & \hat{X}_2 & \hat{X}_3 & \hat{X}_4 \end{bmatrix} \quad (3.9)$$

La relation 3.8 n'est pas tout le temps valide. Toutefois, en l'exprimant sous la forme d'un produit vectoriel, une solution linéaire pour  $\hat{M}_{hom}$  peut être écrite sous la forme :

$$\hat{X}'_i \times \hat{M}_{hom} \cdot \hat{X}_i = 0 \quad (3.10)$$

avec

$$\hat{M}_{hom} \cdot \hat{X}_i = \begin{bmatrix} h^{1T} \\ h^{2T} \\ h^{3T} \end{bmatrix} \cdot \hat{X}_i = \begin{bmatrix} h^{1T} \cdot \hat{X}_i \\ h^{2T} \cdot \hat{X}_i \\ h^{3T} \cdot \hat{X}_i \end{bmatrix} \quad (3.11)$$

et

$$\hat{X}'_i = \begin{pmatrix} x'_i \\ y'_i \\ w'_i \end{pmatrix} \quad (3.12)$$

Le produit vectoriel peut alors s'écrire :

$$\hat{X}'_i \times \hat{M}_{hom} \cdot \hat{X}_i = \begin{bmatrix} 0^T & -w'_i \hat{X}_i^T & y'_i \hat{X}_i^T \\ w'_i \hat{X}_i^T & 0^T & x'_i \hat{X}_i^T \\ -y'_i \hat{X}_i^T & x'_i \hat{X}_i^T & 0^T \end{bmatrix} \begin{pmatrix} h^1 \\ h^2 \\ h^3 \end{pmatrix} = 0 \quad (3.13)$$

L'expression 3.13 prend la forme  $\hat{A}_i \cdot h = 0$  avec  $h$  le vecteur colonne formé des neuf éléments de la matrice  $\hat{M}_{hom}$ . Pour chacun des quatre sommets, on calcule cette matrice  $\hat{A}_i$  afin d'obtenir la matrice  $\hat{A} = [\hat{A}_1 \quad \hat{A}_2 \quad \hat{A}_3 \quad \hat{A}_4]^T$ . L'ensemble d'équations  $\hat{A} \cdot h = 0$  est résolu à l'aide d'une approche des moindres carrés.

## Annexe B : Calcul des critères $C$ et $e$ du module de détermination de posture

Une méthode d'analyse des composantes principales (Principal Component Analysis, PCA) est utilisée pour la détermination de  $C$  et  $e$ . L'idée principale est de déterminer la projection qui représente le mieux les données au sens des moindres carrés [Dud01]. On considère que l'origine des deux axes principaux (on dispose d'une forme en 2D) correspond au point moyen  $C$  de l'ensemble du nuage de points. Soit  $C_N$ , le nuage de  $N$  points issu de la segmentation d'arrière plan. L'objectif est de trouver le vecteur  $e$  passant par l'origine  $C$  de telle sorte que la somme des distances au carré entre les différents points  $p_i \in C_N$  et leurs projections respectives  $p_i^*$  sur  $e$  soit la plus petite possible. On recherche donc à minimiser la fonction suivante :

$$J(p_1^*, \dots, p_N^*, e) = \sum_{i=1}^N \|p_i^* - p_i\|^2 \quad (3.14)$$

En considérant que le vecteur  $e$  est un vecteur unitaire, chaque point  $p_i^*$  peut s'écrire de la manière suivante :

$$p_i^* = C + \alpha_i e \quad (3.15)$$

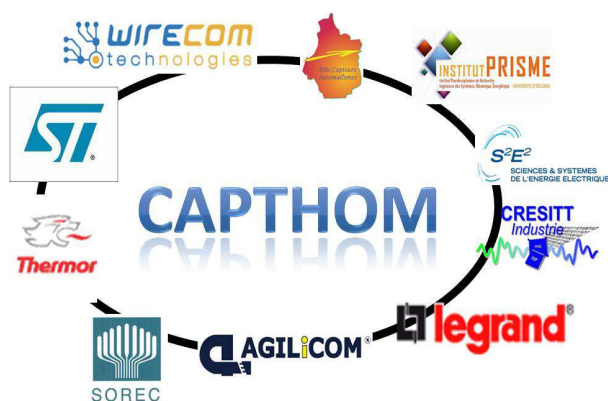
avec  $\alpha_i \in \mathbb{R}$ . Le principe revient donc à déterminer l'ensemble optimal de coefficients  $\alpha_i$  en substituant (3.15) dans (3.14) et en minimisant la nouvelle fonction :

$$J(\alpha_1, \dots, \alpha_N, e) = \sum_{i=1}^N \|(C + \alpha_i e) - p_i\|^2 \quad (3.16)$$

En ajoutant la contrainte  $\|e\| = 1$  et en fixant la dérivée à zéro pour obtenir la meilleur direction  $e$ , la résolution du problème implique la détermination de la matrice de covariance  $A$  définie de la manière suivante :

$$A = \sum_{i=1}^N (p_i - C)(p_i - C)^T \quad (3.17)$$

Le vecteur propre correspondant à la plus grande valeur propre de la matrice de covariance  $A$  est le premier vecteur principal  $e$ .



Nous tenons à remercier les partenaires du projet CAPTHOM. Ce travail a été réalisé avec le soutien financier de la Région Centre et du Ministère de l'Industrie dans le cadre du projet CAPTHOM du pôle  $S^2E^2$ , [www.s2e2.fr](http://www.s2e2.fr).



# Damien BRULIN

## Fusion de données multi-capteurs pour l'habitat intelligent

Le concept d'habitat intelligent s'est largement développé ces dernières années afin de proposer des solutions face à deux préoccupations majeures : la gestion optimisée de l'énergie dans le bâtiment et l'aide au maintien à domicile de personnes âgées. C'est dans ce contexte que le projet CAPTHOM, dans lequel s'inscrit cette thèse, a été développé. Pour répondre à ces problématiques, de nombreux capteurs, de natures différentes, sont utilisés pour la détection de la présence humaine, la détermination de la localisation et de la posture de la personne. En effet, aucun capteur, ne peut, seul, répondre à l'ensemble de ces informations justifiant le développement d'un dispositif multi-capteurs et d'une politique de fusion de données. Dans ce projet, les capteurs retenus sont les détecteurs infrarouges passifs, les thermopiles et la caméra. Aucun capteur n'est porté par la personne (non invasivité du dispositif). Nous proposons une architecture globale du capteur intelligent composée de quatre modules de fusion permettant respectivement de détecter la présence humaine, de localiser en 3D la personne, de déterminer la posture et d'aider à la prise de décision finale selon l'application visée. Le module de détection de présence fusionne les informations des trois capteurs : les détecteurs IRP pour la détection du mouvement, les thermopiles pour la présence en cas d'immobilité de la personne et la caméra pour identifier l'entité détectée. La localisation 3D de la personne est réalisée grâce à l'estimation de position sur horizon glissant. Cette méthode, nommée Visual Receding Horizon Estimation (VRHE), formule le problème d'estimation de position en un problème d'optimisation non linéaire sous contraintes dans le plan image. Le module de fusion pour la détermination de posture s'appuie sur la théorie des ensembles flous. Il assure la détermination de la posture indépendamment de la personne et de sa distance vis à vis de la caméra. Enfin, un module d'aide à la décision fusionne les sorties des différents modules et permet de déclencher des alarmes dans le cas de la surveillance de personnes âgées ou de déclencher des applications domotiques (chauffage, éclairage) pour la gestion énergétique de bâtiments.

Mots clés : Fusion de données, estimation sur horizon glissant, logique floue, habitat intelligent

### Multi-sensors data fusion for smart home

The smart home concept has been widely developed in the last years in order to propose solutions for two main concerns : optimized energy management in building and help for in-home support for elderly people. In this context, the CAPTHOM project, in which this thesis is in line with, has been developed. To respond to these problems, many sensors, of different natures, are used to detect the human presence, to determine the position and the posture of the person. In fact, no sensor can , alone, answers to all information justifying the development of a multi-sensor system and a data fusion method. In this project, the selected sensors are passive infrared sensors (PIR), thermopiles and a video camera. No sensor is carried by the person (non invasive system). We propose a global architecture of intelligent sensor made of four fusion modules allowing respectively to detect the human presence, to locate in 3D the person, to determine the posture and to help to make a decision according to the application. The human presence module fuses information of the three sensors : PIR sensors for the movement, thermopiles for the presence in case of immobility and the camera to identify the detected entity. The 3D localisation of the person is realized thanks to position receding horizon estimation. This method, called Visual Receding Horizon Estimation (VRHE), formulates the position estimation problem into an nonlinear optimisation problem under constraints in the image plane. The fusion module for the posture determination is based on fuzzy logic. It insures the posture determination regardless of the person and the distance from the camera. Finally, the module to make a decision fuses the outputs of the preceding modules and gives the opportunity to launch alarms (elderly people monitoring) or to command home automation devices (lightning, heating) for the energy management of buildings.

Key words : Data fusion, receding horizon estimation, fuzzy logic, smart home



Institut PRISME  
Polytech'Orléans - 8, rue Léonard de Vinci  
45072 Orléans France

